Screening, isolating and characterizing acetyl xylan esterase enzymes from a novel ecological niche

By Charl Marais

Thesis presented in fulfilment of the

requirements for the degree of in the Faculty of Natural Sciences

at Stellenbosch University

Supervisor: Prof. Jens Kossmann

`

Co-Supervisors: Dr. Gavin M George

March 2015

Institute for Plant Biotechnology, Stellenbosch University, Private Bag X1, Matieland, 7602, Stellenbosch, South Africa

December 2014

Declaration

By submitting this thesis/dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

01/03/2015

Summary

Hemicellulose, a plant cell wall polysaccharide composed amongst other 6-carbon sugars of xylose, a 5-carbon fermentable sugar, has various side-chains which hinder breakdown by the hydrolytic enzyme xylanase. We executed a screen of a metagenomic library established from heat treated saw dust to isolate novel genes for plant cell wall esterases and isolated two clones that potentially could be utilised as acetyl xylan esterases: HEL12 and HEL37 from *Citrobacter farmeri* and *Bacillus vallismortis* respectively are confirmed short-chain acetyl xylan esterases (C2-C4) with an optimal temperature of 30°C and 35°C and pH 8.0. HEL12 (33kDa) and HEL37 (25kDa) are small, dextrous acetyl xylan esterases with HEL37: K_m of 1.621mM for p-nitrophenyl acetate and K_m of 3.571mM for the substrate p-nitrophenyl palmitate with a V_{max} of 2.462 mMol/min/mg protein and V_{max} of 0.4363 mMol/min/mg protein respectively. In addition HEL12 was active on other substrates with a Km of 1.321mM protein for p-nitrophenyl acetate and K_m of 1.692mM for the substrate p-nitrophenyl butyrate with a V_{max} of 3.812 and V_{max} of 1.523 mMol/min/mg protein respectively. Both enzymes were assayed on various acetylated polymers including acetylated xylan to indicate their ability to hydrolyse plant lignocellulosic polymers.

Keywords: AE, Acetyl Xylan Esterase, Lypolitic, p-Nitrophenyl-Esters, Library, Tributyrin, Phylogenetic

Opsomming

Met die herwinde interisering in die energie kriesis, is “herwinbare energie” vinnig besig om die nuwe modewoord te raak tussen verskeie in industriële bedrywighede. Hemicellulose, ‘n plant sel wand polisakkaride opgemak uit xilose, ‘n 5-koolstof fermenteerbare suiker, het verskillende sykettings wat ‘n hidernis veroorsaak in die hidrolase van die komponent deur xilanase ensieme. HEL12 en HEL37 van bakterieë *Citrobacter farmeri* en *Bacillus vallismortis* afsonderlik, is deel van die hidrolitiese groep van ensiemme wat die hidrolisasie van ester sy-bindings op die xilaan suikerstring kataliseer. Die ensieme was primêr geassesseër vir aktiwiteit deur middel van sintetiese substraat 4-nitrofeniel asetaat. HEL12 en HEL37 was gevind om kort-ketting asetiel xilaan esterase (C2-C4) te wees, met ‘n optimale temperatuur van 30°C en 35°C afsonderlik, beide met ‘n optimale pH van 8.0. Met ‘n geskatte gewig van omtrent +- 30kDa is beide die ensiemme van die kleiner meer behendige asetiel xilaan esterase, met HEL37: K_m = 1.621 op 4-nitrofeniel asetaat en K_m = 3.571 op substraat 4-nitrofeniel palmetaat, en V_{max} = 2.462 met ‘n V_{max} = 0.4363 afsonderlik, en HEL12: K_m = 1.321 op 4-nitrofeniel asetaat en K_m = 1.692 op sustraat 4-nitrofeniel buteraat met ‘n V_{max} = 3.812 en V_{max} = 1.523 afsonderlik.

Kern Woorde: AXE, Asetiel Xilaan Esterase, Lipolitiese, 4-nitrofeniel, Biblioteek, Filogenetika

Acknowledgements

First and foremost all the glory of life and science goes to GOD. Thank you for giving me the opportunity and mind-set to allow me to complete a research project aimed at protecting your planet. Without your guidance I would have been lost.

Thank you to my family for standing by me, and having an interest and supporting me through this effort. Thanks for the understanding, the guidance and the love that gave me hope and a smile.

The privilege of allowing me to do an MSc in the IPB was extended to me by my supervisor Prof. Jens Kossmann. Thank you for being there to set me on a path of discovery and being there for financial support. I will forever remember the knowledge and wisdom gained from you.

To my co-supervisors, Dr. Gavin George and Dr. Inonge Mulako: There were hard times and there were easy ones, there was frustration, and there was joy… and then there was the pat on my shoulder every now and again. I thank you for consultation, planning and project construction.

To the in-lab specialists, Marnus Smith, Dr. Stanton Hector: It was an incredible ride for me, thanks for sharing intricate knowledge and experience in specialised fields.

Eternal thanks to Mr. Kyle Willard that built this incredible library for the study, and allowed me to pester him with a large amount of questions. You, are a very interesting individual, may your success only be equalled by your dreams.

Special thanks to Dr. Paul Hills, for listening, honesty, and openness. You are someone I will always feel comfortable with talking through problems, and discussing theories.

Dr. Shaun Peters, thank you for all the meaningful discussions. You meant more to me than you probably realise. Thanks for allowing me to get to know you.

I am grateful for the phylogenetic knowledge passed to me by Prof. Karin Jacobs. The knowledge was invaluable to my studies.

Mr. George Fredricks, thank you for making the lab a more colourful place to be in.

A special thanks to the Lüderitz Christian School (ACE), the supervisors and monitors that guided my growing mind. I want this dissertation to stand as a testimonial, that no door is ever shut, one just needs to push harder. You inspired me to become more than I thought possible. Through solid education, and the understanding that my decisions lead to my own success, you helped me build a solid future.

To the Lüderitz community itself and friends, you have shown me that no Buchter ever stands alone. Thank you for the support, the travel opportunities, care, love and most especially advice that each of you have invested in me. They say “It takes a village” I am blessed to say that my “village” has a heart larger than the town itself.

Lastly, thank you to all my friends, fellows, and partners. The conversations were always meaningful, and I long for a time where I will again, have more time to offer you. I would like to thank every student and staff member of the Institute for Plant Biotechnology. Each of you gave me insight and knowledge that I would trade for nothing.

Table of Contents

Chapter 1

1. Introduction

Fossil fuels have been the major source of energy since the early twentieth century (Shafiee and Topal, 2009). For example, The Institute for Energy Research, based in Washington D.C., USA, reported that fossil fuels constitute 95% of the energy used in the transport sector in the USA (http://www.instituteforenergyresearch.org/energy-overview/fossil-fuels/). Furthermore, there is a growing demand for energy due to the expanding human population and increased industrialization (Fontana *et al.,* 2013; Pérez-Lombard *et al.,* 2008). The human population increased 270% (2.5 to 6.8 billion) from 1950 to 2009 (http://www.un.org). This increased demand in energy consumption has led to a decline in global fossil fuel reserves (Shafiee and Topal, 2009). As a consequence of the increase in human population, increase in energy demand and the decrease in fossil fuel reserves, the development of alternative renewable energy sources is of critical importance (Demirbas, 2006). Renewable energy is defined as any naturally-occurring energy source that is rapidly replaced by nature and is therefore inexhaustible (Barbier, 2002). There are different kinds of renewable energy sources (Figure 1.1). These include; nuclear power, solar energy (Twidell and Weir, 2006), geothermal energy (Lund and Freeson, 2001), wind (Grauers, 1996) and biomass (Kim and Dale, 2006).

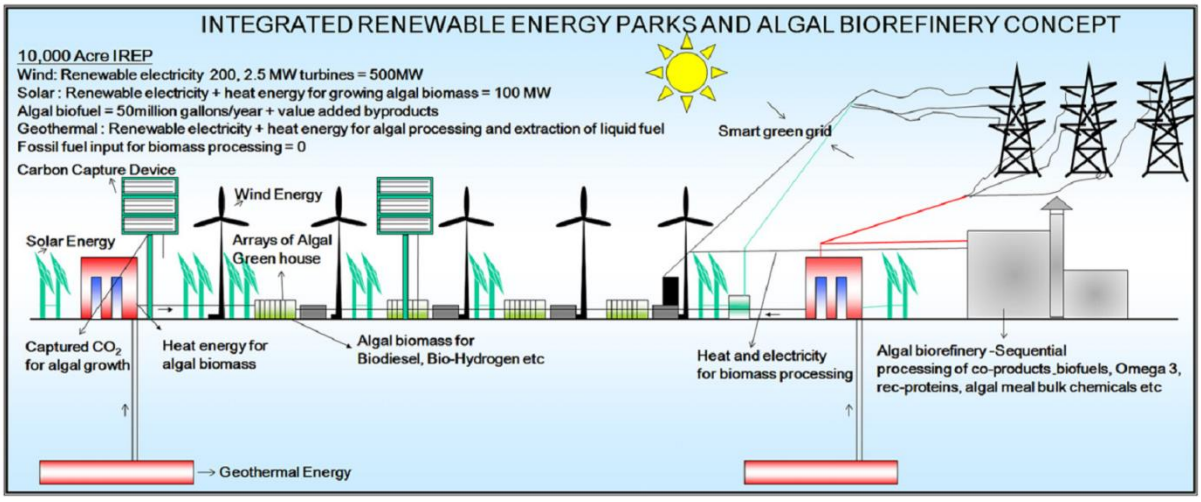

Figure 1.1. A concept for a renewable energy farm showing bioprocessing, wind energy, solar energy, thermal and geothermal energy (Subhadra, 2010).

Biofuel is fuel created from renewable biomass that is harvested in a sustainable way, or biological waste from another industrial process (Reijnders, 2010). Biomass can be utilized to generate biodiesel from the complex oils found in the seeds of sunflower, canola and jatropha (Ragaglini *et al.,* 2011; Berchmans and Hirata, 2008; Antolin *et al.,* 2002). Biofuel can also be generated by means of fermentation of sugar complexes (Reijnders, 2010). The enzymatic breakdown of biomass into simple sugars for fermentation into bio-ethanol and bio-butanol production is a common focus of research (Pauly and Keegstra, 2010). The most common and highly researched of these biofuels is bio-ethanol. Bio-ethanol production methods can be sub-divided into two main forms; first and second generation bio-ethanol production.

1.1 First generation bio-ethanol production

First generation bio-ethanol production relies on biomass with a high starch or sucrose content (Naik *et al.,* 2010). The countries with the highest production rate of ethanol by these means are Brazil and USA, where production utilises either corn or sugarcane (USDA FAS, 2014). In Europe the most common forms of biomass for ethanol production are potato, wheat and sugar beet (USDA FAS, 2008). The biomass is treated with heat and enzymatic digestion (Figure 1.2) to release simple sugars such as glucose and fructose for fermentation by the baker's yeast *Saccharomyces cerevisae* into ethanol (Clay, 2003).

This type of bio-ethanol production impinges directly on the food industry, since starch is a valuable food resource (Naik *et al.,* 2010). Other than competing directly with the food market, serious doubts have been raised as to the ethical reasoning of using a possible food source for fermentation and production of what is seen by many as a commodity (Naik *et al.,* 2010). When new ground is developed to produce a surplus of this crop, specifically for biofuel production, the net loss in land use, deforestation and indirect land use with regards to water usage often becomes a problem (Havlik *et al.,* 2011). Plant material that is not consumable by humans is of the greatest advantage, since it would negate the ethical dilemma of food plantations in a starving world (Fraiture *et al.,* 2007). This is known as second generation biofuel production (Naik *et al.,* 2010; Antizar-Ladislao and Turron-Gomez, 2008; Schenk *et al.,* 2008).

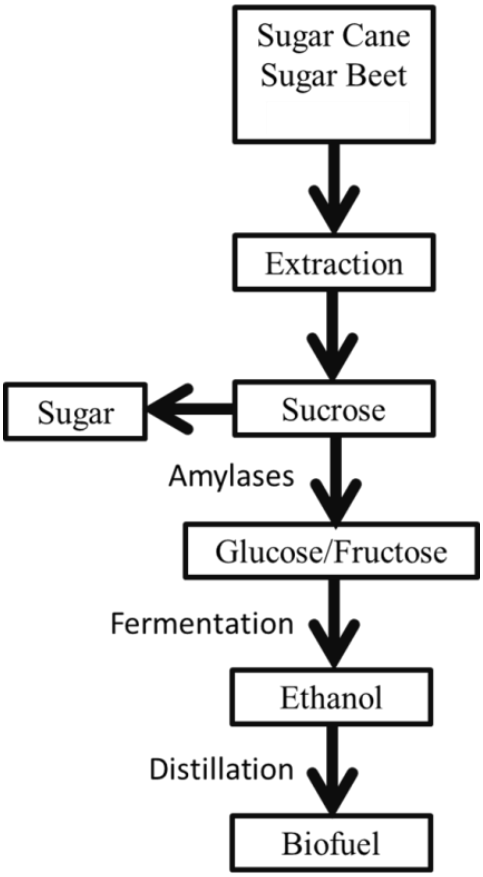

Figure 1.2: Depiction of first generation Biofuel production. The substrate, sugar cane, sugar beet, or starch is extracted and broken down to sucrose, which is treated by enzymes to break it down to monomeric sugars, glucose and fructose. Both of these sugars are fermented by *Saccaromyces cerevisae* and other yeasts. The ethanol is heat-distilled to a concentration that is combustible as a biofuel. (adapted from Caffall and Mohnen, 2009)

1.2 Second generation bio-ethanol production

The focus of second generation biofuel technology is on the breakdown of non-food plant material into simple sugars for fermentation to produce ethanol or butanol (Naik *et al.,* 2010; Antizar-Ladislao and Turron-Gomez, 2008; Schenk *et al.,* 2008). The aim is to use sugarcane bagasse, maize stalks and fast growing green crops like switchgrass (Figure 1.2) (Prasifka *et al.,* 2010). The disadvantage of this is the complex structure of the plant cell wall which consists of cellulose, hemicellulose and lignin that are difficult to break down to simple sugars (Naik *et al.,* 2010; Antizar-Ladislao and Turron-Gomez, 2008; Schenk *et al.,* 2008). The sources of biomass listed in Figure 1.2 are rich in lignocellulose (Abramson *et al.,* 2010). Lignocellulose is broken down by various enzymes and chemical processes (Figure 1.3), which is currently time consuming and expensive (Sanderson, 2011), consequently it has not yet been fully established for biofuel production.

Figure 1.3. Examples of substrates used for second generation bio-ethanol production. (a) bagasse from *Saccharum* species (Bahcegul *et al.,* 2012), (b) *Zea mays* cane (Kapanigowda *et al.,* 2010), (c) Switchgrass (Prasifka *et al.,* 2010)

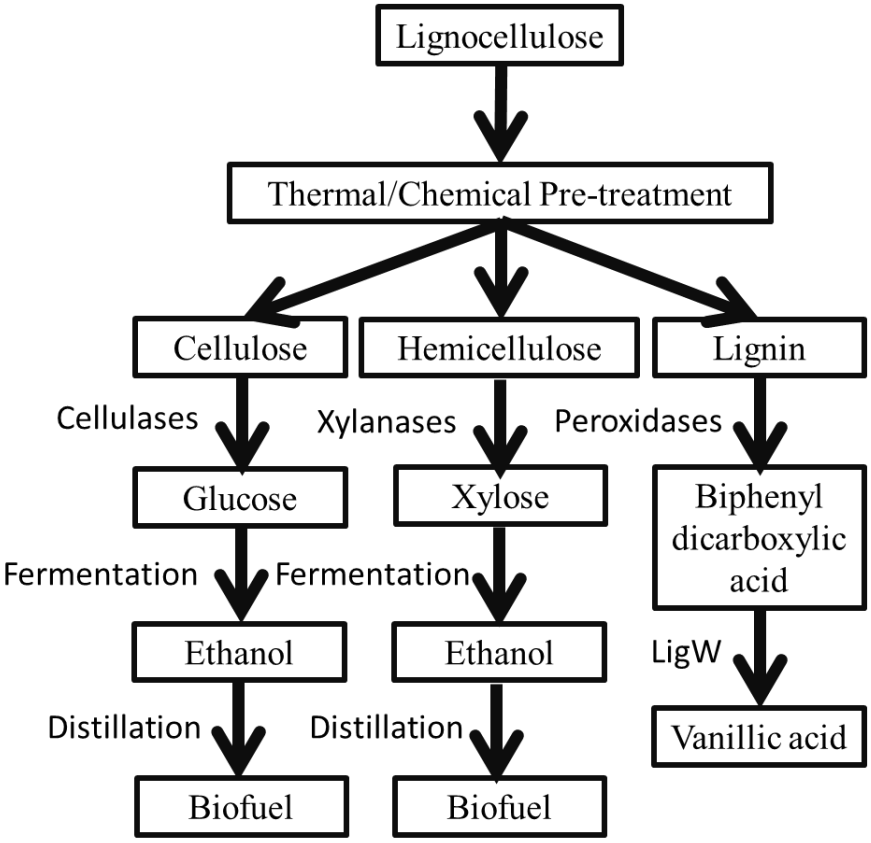

Figure 1.3. Depiction of second generation biofuel production. The substrate, lignocellulose (plant cell wall) is treated and denatured via thermochemical means to be able to separate cellulose, hemicellulose and lignin complexes. Each individual component is treated by substrate specific hydrolases, which break down the components into less complex monomers like glucose, xylan and biphenyl-dicarboxylic acid. The monomeric sugar compounds are fermented by *Saccaromyces cerevisae*, and the biphenyl-dicarboxylic acid by LigW enzymes found in the bacterial biome in the digestive tract of lignocellulose degrading insects. The fermentation product is ethanol, which can then be distilled to a combustible concentration for use as a biofuel (adapted from Caffall and Mohnen, 2009).

1.3 Composition of Lignocellulose

Lignocellulose is composed mainly of four polymers; cellulose, lignin, pectin and hemicellulose (Keegstra, 2010; Caffall and Mohnen, 2009; De Vries and Visser, 2001). These constituents of lignocellulose are amalgamated by a complex collection of interconnecting moieties (Sorek *et al.,* 2014). Plant cell wall polysaccharides are an abundant and renewable organic carbon source as they are mainly composed of fermentable carbohydrate monomers (Abramson *et al.,* 2010; Caffall and Mohnen, 2009). The deconstruction of this polymer into its monomeric units for fermentation is achieved either by a complex array of enzymes or chemicals working in a cohort to completely hydrolyse the plant cell wall (Schimpf *et al.,* 2013; Verma *et al.,* 2013). These include, but are not limited to endoglucanases, releasing glucooligosaccharides from cellulose hydrolization, cellobiohydrolases which hydrolyse crystalline cellulose to form cellubiose, β-glucosidases and exoglucanases which further hydrolyse the glucooligosaccharides to glucose, endoxylanases hydrolysing the hemicellulosic xylan backbone and esterases which specialize in the removal of esterified side chains on various sugar backbones (Sanderson, 2011; De Vries and Visser, 2001). Acetyl xylan esterases hydrolyse and remove acetyl groups attached to the O-2 or O-3 position on the xylan backbone of hemicellulose (Verma *et al.,* 2013; Biely *et al*., 1991; Khan *et al.,* 1990).

1.3.1 Cellulose

The most abundant polysaccharide, constituting 40-60% of lignocellulose, is cellulose, which is composed of glucose linked with β-1-4 glycosidic bonds (O'Sullivan, 1996). This makes for a linear polymeric structure with no side groups attached. The simple nature of this polymer in lignocellulose makes it the ideal substrate for biofuel production (Klemm *et al.,* 2005). Cellulose, being so uniform in nature (Figure 1.4), is crystalline and therefore not readily amenable to enzymatic attack (Igarashi *et al.,* 2011; Park *et al.,* 2010; Tian *et al.,* 2009). Developments have highlighted methods of increasing the diversity of cellulases and glycosyl hydrolases that hydrolyse cellulose to cellobiose and glucose via enzymatic hydrolysis (Henrissat and Bairoch, 1993; 1996; Xiros *et al.,* 2013). Most isolated cellulases are derived from fungi, most notably from *Aspergillus niger*, due to the demand for enzymes that are functional at higher temperatures, new research has focussed on thermophyllic bacteria like *Clostridium thermocellum* (Oberoi *et al.,* 2014; Lv and Yu 2012).

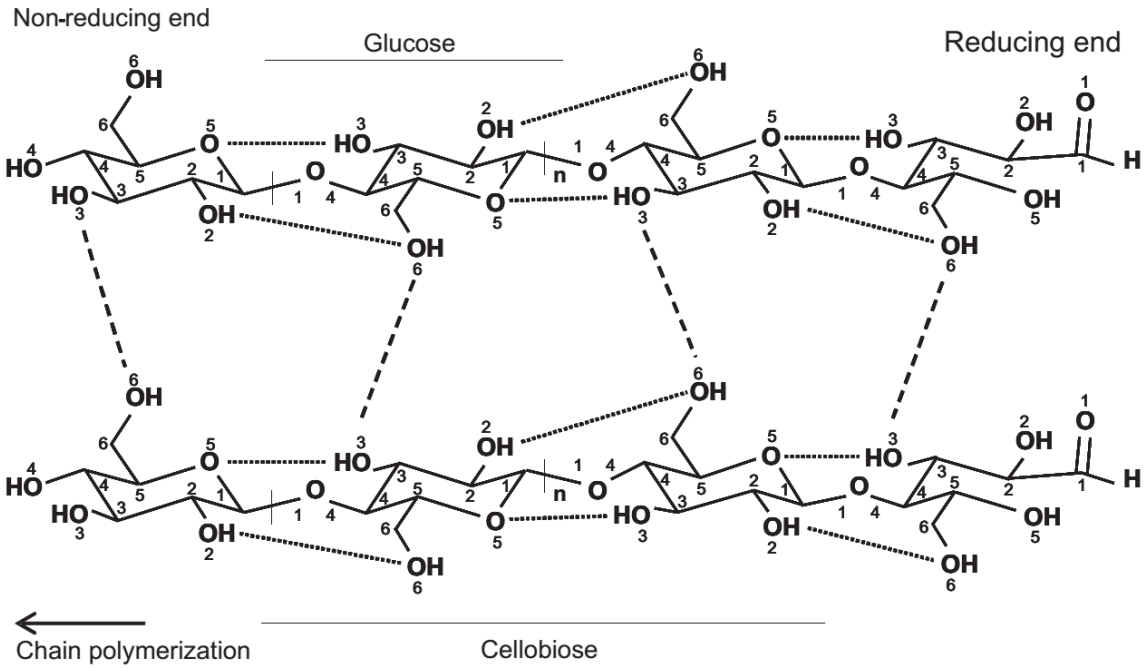

Figure 1.4. Cellulose structure showing the inter- and intra-chain alpha hydrogen bonds. Lines with dashes indicate bonds between separate chains and dotted lines indicate hydrogen bonds that form within the same chain (Chen *et al.,* 2012).

1.3.2 Pectin

Comprising as much as 30% of the angiosperm cell wall, pectin is a heteropolysaccharide with polygalacturonic acid substituted with rhamnose as a backbone (Chen *et al.,* 2012; Pavcová and Hacke, 2011; Caffall and Mohen, 2009). The pectin backbone can have various side chains, for example ferulic acid, diferulic acid and uronic acid (Ridley *et al.,* 2001; Willats *et al.,* 1999; De Vries *et al.,* 1982). This configuration of the complex polysaccharide is important to the cell wall due to the structure it forms by cross-linking between hemicellulose, phenolic compounds and proteins within the plant cell wall (Popper and Fry, 2005; Nakamura *et al.,* 2002; Bauer *et al.,* 1973). Most of the commonly used commercial enzymes that break down pectin are produced in filamentous fungi, for example, *Botrytis cinerea* (Benoit *et al.,* 2012; Shah *et al.,* 2009). The enzymes used in pectin deconstruction are glycoside hydrolases, polysaccharide lyases, carbohydrate esterases and pectases (Benoit *et al.,* 2012; Shah *et al.,* 2009).

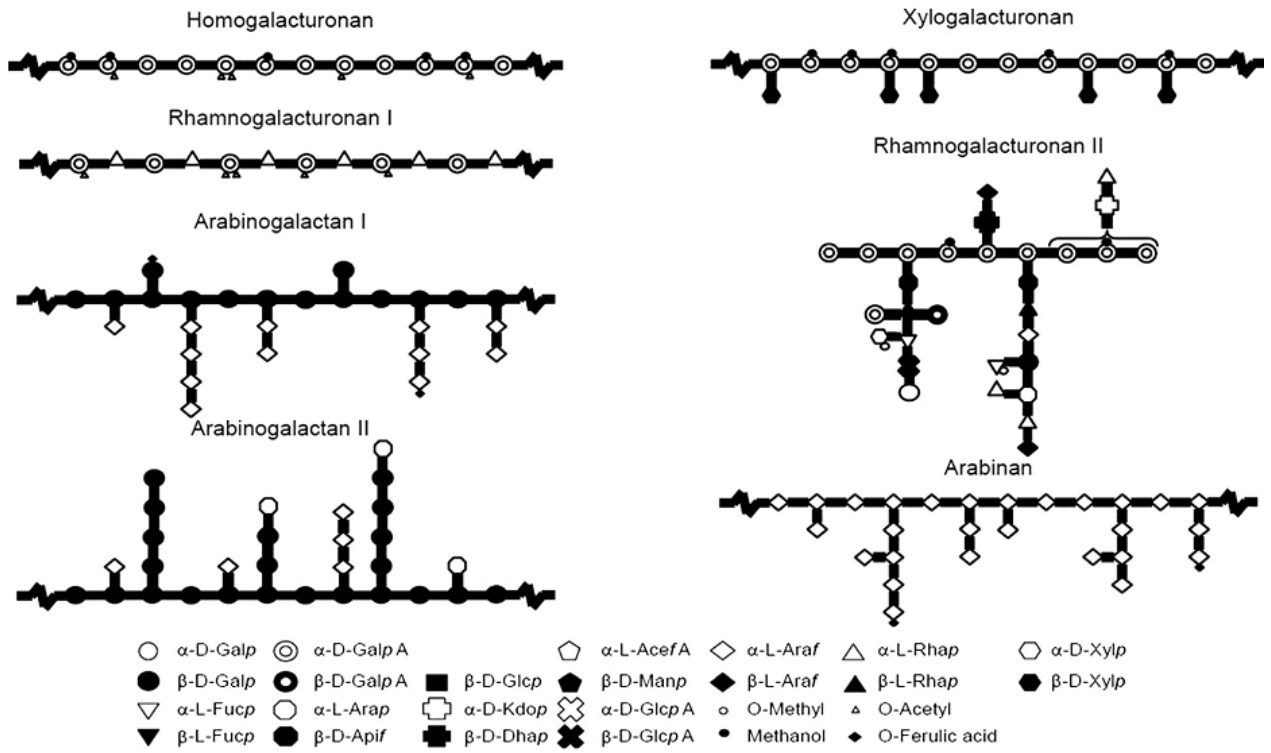

Figure 1.5. The structure of various pectin polymers depicting their side-chains and polysaccharide constituents (Benoit *et al.,* 2012).

1.3.3 Lignin

Of the aromatic compounds that exist in lignocellulose, lignin is the most prominent, constituting between 10-30% of the lignocellulose, cross-linking the different homo- and hetero-polysaccharides in the higher plant cell wall (Bugg *et al.,* 2011). Lignin (Figure 1.6) is a heteropolymer consisting of phenylpropanoid aryl-C_3 base, consisting of three different alcohols, *trans-p*-coumaryl alcohol, *trans-p*-coniferyl alcohol and *trans-p*-sinapyl alcohol (Vanholme *et al.,* 2010; Boerjan *et al.,* 2003; Zaldivar *et al.,* 2001). These alcohols are all methoxylated to different degrees forming the phenylpropanoid (Boerjan *et al.,* 2003). The degree of natural polymeration of this hydrophobic aromatic compound is difficult to determine since it disintegrates when purified (Vanholme *et al.,* 2010). Purified lignin can become as large as 10000 monomeric units long, and have a vast number or methoxylated alcohols (Vanholme *et al.,* 2010). Lignin deposits are mostly found on the cell walls of secondary thickened cells, allowing them to be rigid and unyielding to stress (Boerjan *et al.,* 2003; Ralph *et al.,* 2004). It is found to be situated between cellulose chains, hemicellulose linking microfibrils and pectin components of cell walls, especially in tracheid, sclereid and xylem cells (Boerjan *et al.,* 2003). The cross-linking makes it impossible to utilize the polysaccharides in fermentation directly, and lignocellulose therefore needs to undergo

various chemical, heat and enzymatic pre-treatments to be able to enzymatically release the simple sugars used for fermentation (Znameroski *et al.,* 2012; Zaldivar *et al.,* 2001). Bacteria that have shown the ability to break down lignin have been found mostly in the intestinal tract of termites and notably include *Streptomyces sp, Acinetobacter sp* and *Arthrobacter globiformis* (Bugg *et al.,* 2011)

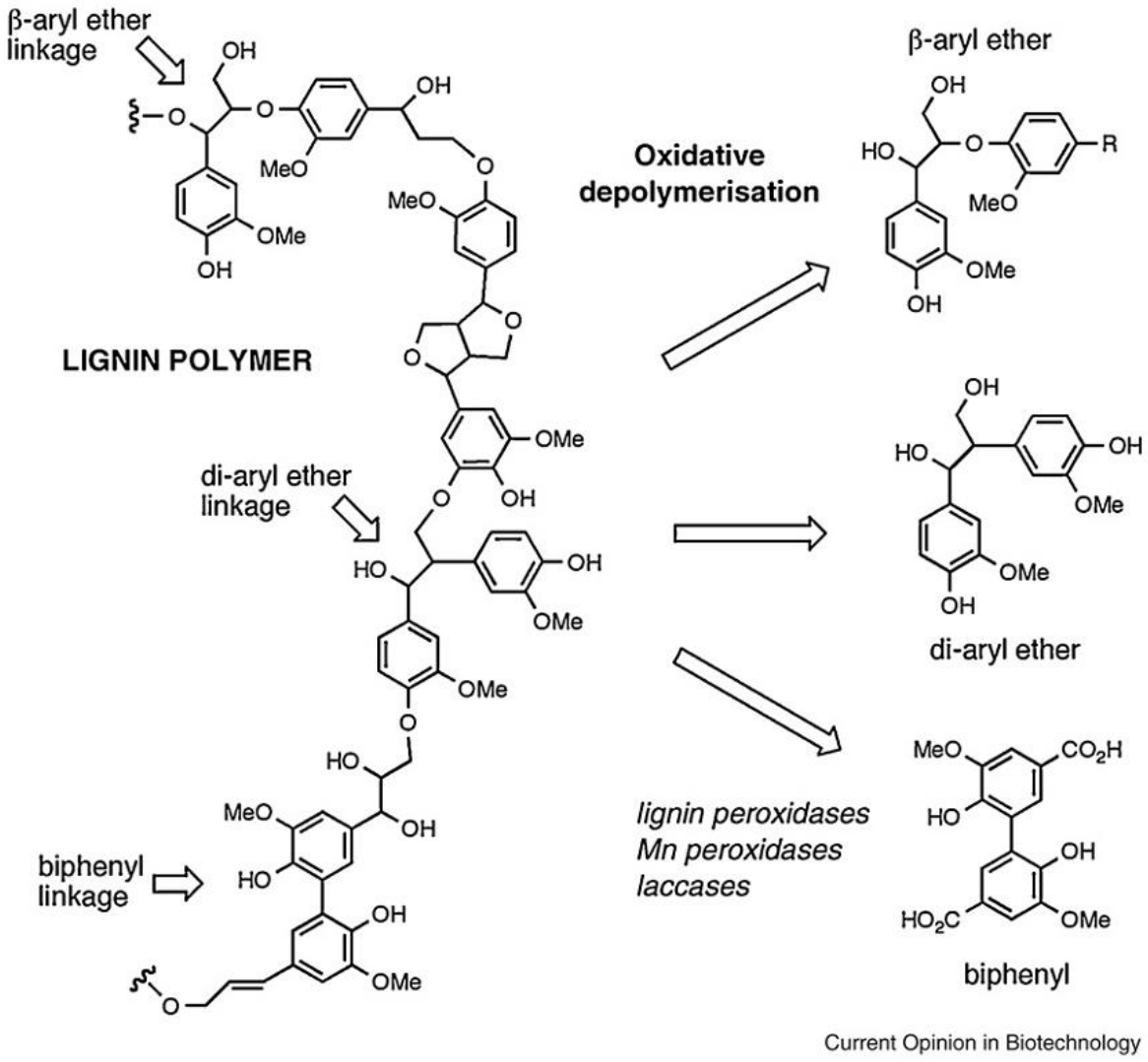

Figure 1.6. The composition and depolymerisation of lignin by chemical means (Bugg *et al.,* 2011).

1.3.4 Hemicellulose

In higher plant-derived biomass esterified bonds are often found on the hemicellulose with a xylan backbone and are subject to hydrolysis by AE enzymes. Recently hydrolysis of hemicellulose polysaccharides has gained attention as an alternative renewable energy source due to rising fossil fuel costs and the green movement. Xylan being a prominent

polymer in hemicellulose of the different softwood plants creates opportunity to hydrolyse this polymer into simple sugars for fermentation into an ethanol product.

The second most abundant, yet rarely utilised polysaccharide in the cell wall is hemicellulose (Figure 1.7). Ranging from 30-50% of the cell wall and being more heterogenous in nature than cellulose, the main constituent of hemicellulose in angiosperms is a heteroxylan polymer consisting of β-1-4-linked D-xylose and L-arabinose, L-rhamnose, D-glucose, D-fructose or D-galactose (Elleuch *et al.,* 2011; De Vries and Visser, 2001; Mok and Antal, 1992). Additions to these heteropolymers can come in the form of D–linked mannuronic acid and galacturonic acid linked covalently to the hemicellulose matrix in angiosperms (Table 1.1). This is common in *Musa* (Banana) species (Jandas *et al.,*2013,; Yang *et al.,* 2007). Hardwoods and cereal lignocellulose contains hemicellulose with an abundance of acetyl substitutions (Gordon and Lomax, 1983). Acetylated hemicellulose forms the focal point for this study.

The linkages of xylan are mosty β-1,4 glycosidic bonds but can also be β-1,6 xylan linkages (Khan *et al.,* 1990). Hemicellulose is synthesised in the Golgi apparatus and transported to the cell wall by vesicles (Mok and Antal, 1992). Hemicellulose chains contain 400-3000 xylose units, as opposed to the 700-15000 glucose units found in the cellulose structure (Saha, 2003). This matrix co-exists and even links to different methyl-esterified pectic polysaccharides, mostly homogalacturonan (Mohnen, 2008). Hemicellulose is known to have side chains attached to the xylan backbone; these can be arabinose, 4-O-methylglucopyranosyl acid, ferulic acid, or an acetyl group (Toth *et al.,* 2013; Margolles-Clark et al. 2009). Most hardwood and annual plants contain xylan polymers with acetyl substitutions at the C2 or C3 position (Peng *et al.,* 2011; Lindenburg et al. 1973, Maekawa 1976). Hemicellulose forms straight microfibrils with high tensile strength and is therefore the perfect scaffold for the plant cell wall (Nishiyama *et al.,* 2002; Brown, 1999). The enzymes that break down hemicellulose are endo-xylanases and β-xylosidases (Kim *et al.,* 2010). The hydrolysis of hemicellulose is hampered by the side chains attached to the xylan backbone (Biely *et al.,* 2014; Badal, 2003). The array of enzymes that remove side chain constituents includes galactosidases, arabinosidases, acetyl esterases, ferulic acid esterases, *p*-coumaroylases, and glucaronic acid esterases (Badal, 2003; Blum *et al.,*1999). These enzymes work together to allow xylanases and xylosidases to hydrolyse xylan into xylose for fermentation (Ademark *et al.,* 1998).

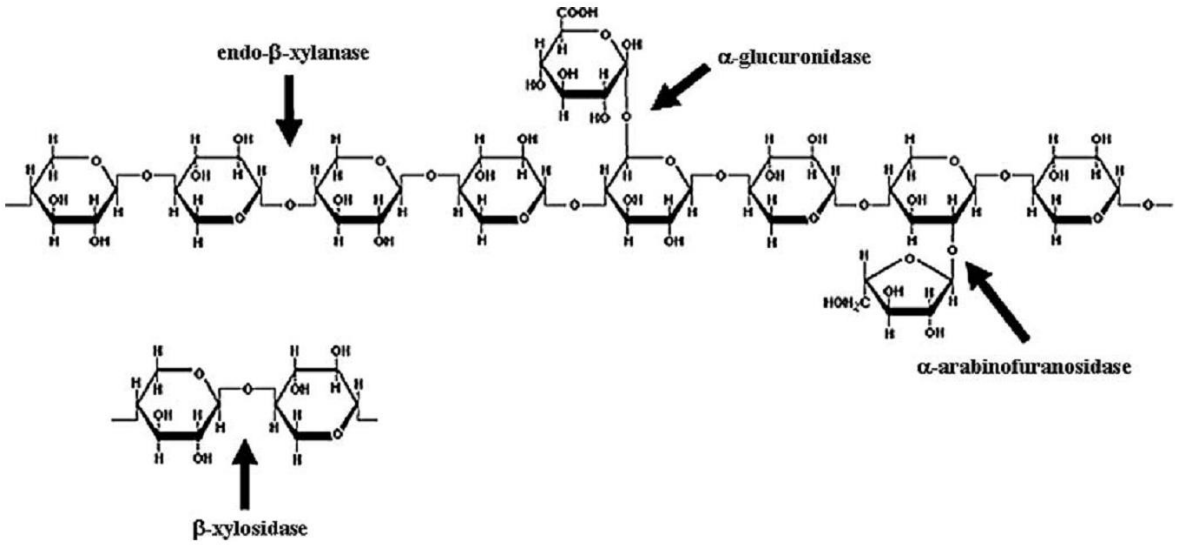

Figure 1.7. Hemicellulose with a β-glucuronic acid residue, and an α-arabinofurane side group attached. The enzymes endo-β-xylanase, β-glucuronidase and α-arabinofuranosidase break the alpha bonds between the subunits and backbone of hemicellulose (Kumar *et al.,* 2008).

Within the last decade many advances have taken place in the industrial degradation of plant cell walls (Zhang *et al.,* 2012). Many of these advancements relate to the enzymatic release of hindering structural groups linked to hemicellulose and lignin (Andric *et al.,* 2010a; Xiao *et al.,* 2004). It has come to a point where the enzymatic pre-treatment and biomass degradation allows access to these biomass resources previously discarded (Khunrong *et al.,* 2011; Rollin *et al.,* 2011; Liu, 2010; Zhang *et al.,* 2010a; Lau *et al.,* 2009). In conjunction with this development, the discovery of novel cellulase and hemicellulase enzymes has enhanced the release of monomeric and short chain sugars from pre-treated biomass (Banerjee *et al.,* 2010a, 2010b, 2010c; Zhang *et al.,* 2010b; Rosgaard *et al.,* 2007; Irwin *et al.,* 1993; Walker *et al.,* 1993). The feasibility of renewable energy sources from green waste relies on further scientific advances in digestion biochemistry of the plant cell wall (Banerjee *et al.,* 2010a; Merino and Cherry, 2007).

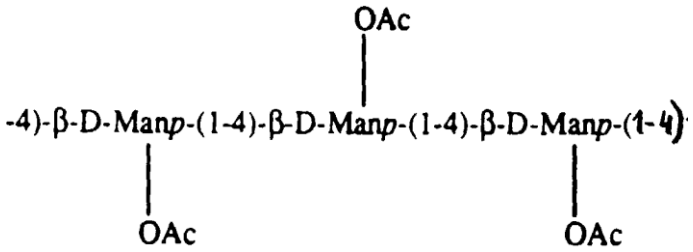

Figure 1.8. Chemical structure of Hemicellulose with a Mannose backbone (Mannan). Mannans are known to be acetylated and the hydrolases that deacetylates O-acetyl mannan groups are acetyl xylan mannanases. The backbone of Mannan can be hydrolysed with mannanases to mannose for fermentation. (http:patentimages.storage.googleapis.com/EP0611304B1/00130001.png)

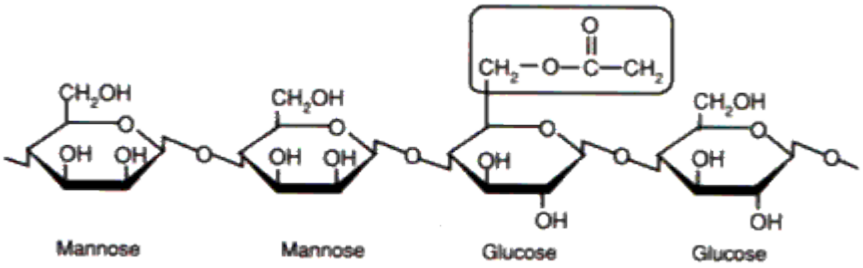

Figure 1.9. Structure of acetylated glucomannan. Mannose is used as an initiating sugar for the linear polysaccharide, with only slight branching. O-Acetyl groups are a commonality with this polymer found in the hemicellulose constituents of the genus *Conifera*. (http:www.konjacfoods.com/images/molecule2.gif)

Mannan and glucomannan, being a prominent polysaccharide of hemicellulose in the plant cell wall. It can consist of a linear or branched polymer polymerised from as D-mannose, D-galactose, and D-glucose. The main component of hemicellulose in softwoods is a glucomannan polysaccharide. Structural studies on the galactosyl side chain indicate that it reacts with the mannan backbone intramolecularly providing structural stability. Acetyl groups are distributed irregularly in glucomannan, making measurement of each of the mannan acetylation assay important each regard. Endo-mannanase and exomannanases facilitates hydolysis of the mannan backbone to oligosaccharides and fermentable sugars. These enzymes include β-mannanase, β-glucosidase, and β-mannosidase. Additional enzymes such as acetyl mannan esterase are required to remove O-acetyl substituents that are incorporated on mannan, thus freeing more of the backbone for enzymatic hydrolysis. These enzymes are of industrial importance to pharmaceutical, food, feed, and pulp and paper industries (Moreira and Filho, 2008). It is shown that Xylanaese and mannanases

hydrolyse a limited number of glyosidic bonds when esterases are not incorporated into the enzyme cocktail (Biely *et al.,*1991).

Sequencing of two Acetyl xylan esterases in *Pseudomonas flourescens* and *Streptomyces lividans* indicated that the binding structure of these enzymes have a catalytic domain that is separate from the binding domain (Margolles-Clark *et al.,* 1996). While the AXE from *Trichoderma reesei* is the most studied enzyme in the field, showed a binding domain similar to fungal cutinases (Margolles-Clark *et al.,* 1996).

Table 1.1: The chemical composition of common dietary fibres (Elleuch *et al.*,2011)

Fibres	Main chain	Branch units	Reference
Cellulose	β-(1,4) glucose		Olson *et al.,* (1987)
β-glucans	β-(1,4) glucose and b-(1,3) glucose		Johansson *et al.,* (2000)
Hemicelluloses			Olson *et al.,* (1987)
Xylans	β-D-(1,4) xylose		
Arabinoxylans	β-D-(1,4) xylose	Arabinose	
Mannans	β-D-(1,4) mannose		
Glucomans	β-D-(1,4) mannose and b-D-(1,4) glucose		
Galactoglucomannans	β-D-(1,4) mannose, b-D-(1,4) glucose	Galactose	
Galactomannans	β-(1,4) mannose	α-D-galactose	
Xyloglucans	β-D-(1,4) glucose	α-D-xylose	
Pectin			
Homogalacturonan	α-(1,4)-D-galacturonic acid (some of the carboxyl groups are methyl esterified)		Ridley *et al.,* (2001)
Rhamnogalacturonan-I	(1,4) galacturonic acid, (1,2) rhamnose and 1-, 2-, 4- rhamnose	Galactose, arabinose, xylose, rhamnose, galacturonic acid	Oechslin *et al.,* (2003)
Rhamnogalacturonan-II		α-(1,4) galacturonic acid	Vidal et al., (2000)
Arabinanes	α-(1,5)-L-arabinofuranose	α-arabinose	
Galactanes	β-(1,4)-D-galactopyranose		
Arabinogalactanes-I	β-(1,4)-D-galactopyranose	α-arabinose	
Arabinogalactanes-II	β-(1,3)- and b-(1-6)-D-galactopyranose	α-arabinose	
Xylogalacturonan	a-(1,4) galacturonic acid	xylose	Le Goff *et al.,* (2001)
Inulin	β-(2,1)-D-fructosyl-fructose		Blecker *et al.,* (2001)
Carrageenan	Sulfato-galactose		
Alginate	β-(1,4)-D-mannuronic acid or α-(1-4)-L-gluceronic acid		
Oligofructose	β-(2,1)-D-fructosyl-fructose		
Polydextrose	D-Glucose		
maltodextrins	α-(1,4)-D-Glucose	α(1,6)-D-Glucose	
Lignin	Polyphenols: Syringyl alcohol, Guaiacyl alcohol, and 4-coumaryl alcohol		Sun *et al.,* (1999)
Chitosan	Polyphenolics: Syringyl alcohol, Guaiacyl alcohol, and 4-coumaryl alcohol, β-(1,4)-linked D-glucosamine and N-acetyl-D-glucosamine		Borderías *et al.,* (2005)

1.4 Plant cell protein targeting peptides

Proteins have specific functions; this often requires a protein to be localized in specific sub-cellular compartments within the cell. Most proteins are synthesised in the cytosol and need to be chaperoned to their designated compartments (Sjöling and Glaser, 1998). This is achieved by addition of a signal targeting peptide to the protein which is then bound to a chaperone protein to relay the protein to its intended destination (Figure 1.10) (Emanuelsson *et al.,* 2000). Not many of the targeting peptides in plants have been elucidated, though they play a large role in the localization of proteins to a specific compartment in the plant cell (Emanuelsson *et al.,* 2007; Raikhel, 1992). Targeting peptides are generally small in nature, being between 17 and 40 amino acids long (Pagny *et al.,* 2003; Yamaguchi *et al.,* 2002). These peptides have been used to create various fusion proteins to allow organelle specific localization of transgenic proteins (Sjöling and Glaser, 1998; Köler *et al.,* 1996 Creissen *et al.,* 1995; Van den Broeck *et al.,* 1985). The targeting of foreign proteins to the chloroplast has been shown to occur when fused to the signal peptide of the small subunit of *Nicotiana tabaccum* ribulose 1.5-bisphosphate carboxylase protein or the *Nicotiana plumbaginifolia* glutathione reductase transit peptide (Creissen *et al.,* 1995; Van den Broeck *et al.,* 1985). Mitochondrial targeting peptides from the cytochrome oxidase subunit IV from yeast at the N-terminus of the protein was successfully used to translocate the green fluorescent protein to the mitochondria of *Arabidopsis thaliana.* This peptide was then cleaved from the protein after it was imported into the mitochondria (Sjöling and Glaser., 1998). A nuclear localization targeting peptide from *Agrobacterium tumefaciens* VirD2 protein was successfully fused to the LacZ protein and targeted the protein to the nucleus of *Nicotiana tabacum* (Herrera-Estrella *et al.,* 1990). This has created a large amount of knowledge in vertebrate targeting peptides. The knowledge regarding protein targeting in plants is limited and therefore most plant signalling peptides are not well known.

Membrane bound proteins in *A. thaliana* often have *N*-linked oligosaccharides (Lerouge *et al.,* 1998). The *N*-glycan oligosaccharides consist of xylose, fructose, glucose and mannose and are added to the oligopeptides by the endoplasmic reticulum and distal Golgi apparatus by glycosyltransferases and xylosyltransferases (Faye *et al.,* 1993; Van Ree *et al.,* 2000). It was shown that glycosyltransferases and xylosyltransferases share an N-terminal transmembrane peptide chain of 17-22 amino acids (Kleene and Berger, 1993; Breton et al., 2001). This transmembrane chain is responsible for these enzymes' localization to the Golgi apparatus by protein-protein interactions or by physical means (Qian *et al.,* 2001; Munro,

1995; Nilsson *et al.,* 1993). The β-1,2-xylosyltransferase protein in *A. thaliana* is no exception to this, as it has a 36 amino acid N-terminal sequence that has been shown to be a cytoplasmic and transmembrane regions of the Xyl-T protein (Pagny *et al.,* 2003). This 36 amino acid sequence, when deleted, inhibited xylosyltransferase activity in *A. thaliana* (Bencúr *et al.,* 2005; Pagny *et al.,* 2003). A fusion of the first 36 amino acids of the *A. thaliana* β-1,2-xylosyltransferase (Xyl-T36) to Green fluorescent protein (GFP) showed localization of this fusion protein to the distal Golgi apparatus (Pagny *et al.,* 2003). Retention of this reporter gene to the Golgi apparatus proves that the Xyl-T36 cytoplasmic and transmembrane domain is responsible for transmembrane association with the Golgi apparatus (Bencúr *et al.,* 2005).

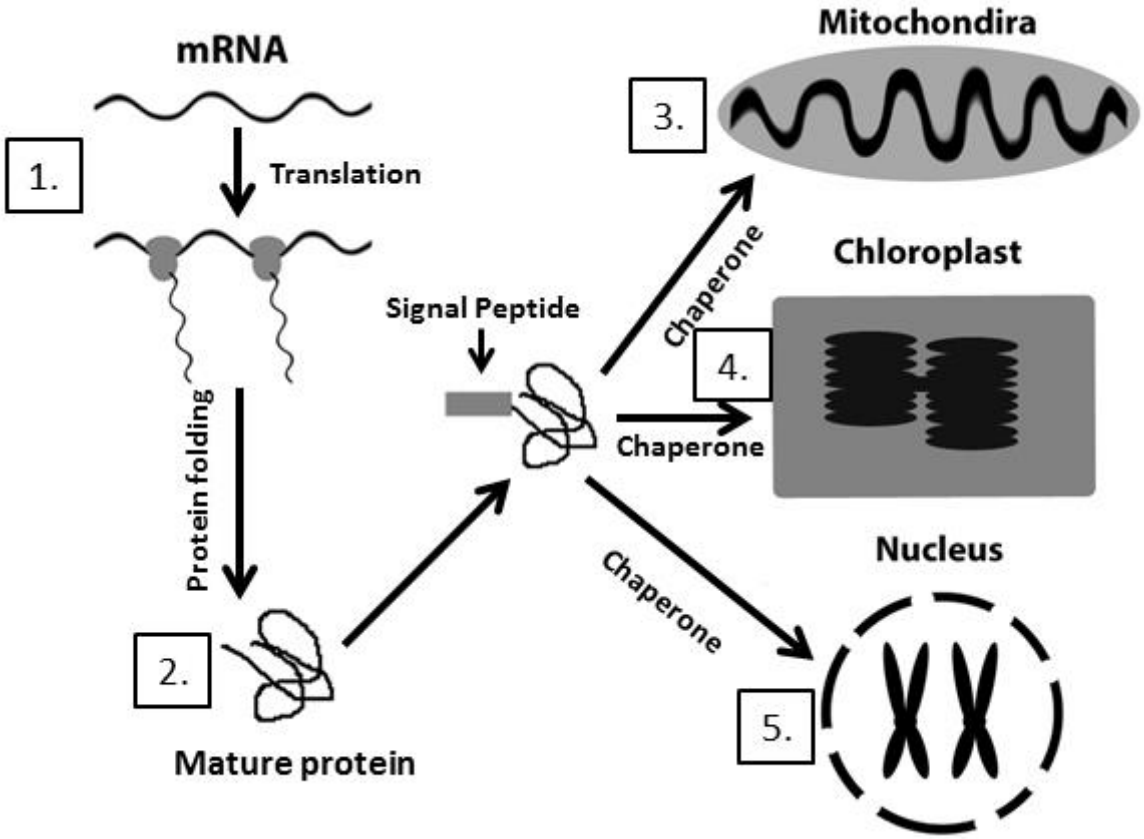

Figure 1.10. Protein transport within the cell. (1.) Protein translation from messenger RNA by ribosomal RNA. (2.) Protein folding after translation. (3.) Protein with cytochrome oxidase subunit IV signal peptide being transported by a chaperone protein to the plant mitochondria (Sjöling and Glaser, 1998). (4.) Protein with small subunit of *Nicotiana tabaccum* ribulose 1.5-bisphosphate carboxylase signal peptide being transported by a chaperone protein to the chloroplast (Creissen *et al.,* 1995; Van den Broeck *et al.,* 1985). (5.) Protein with *Agrobacterium tumefaciens* VirD2 protein signal peptide attached being localised to the nucleus by a chaperone protein (Herrera-Estrella *et al.,* 1990).

Aim of this study

The aim of this research was to identify novel acetyl xylan esterase enzymes.

Objectives of this study

- Therefore we chose to isolate bacteria from a novel niche (Cape Saw mill, Stellenbosch, South Africa),
    - identify them,
    - construct a multi-genomic library,
    - Screen this library for clones conferring acetyl esterase activity,
    - Cloned fragments that express enzymes that screened positive for lipolytic and then secondly for p-nitrophenyl acetyl esterase activity
- Clones were sequenced and possible open reading frames were identified via bioinformatics
- ORF's were then amplified for insertion into the expression vector system pRSET C
- fusion protein for further biochemical characterization.

Chapter 2

2. Materials and Methods

2.1 Sample collection and enrichment

A saw dust sample of 200 g was collected from an autoclave in a Saw Mill (Cape Saw Mills, Stellenbosch, South Africa). The sample was enriched for cell wall degrading thermophilic bacteria by heating to 65°C for 25 days in the presence of 800 ml water and 20 g of pine wood chips. Six dilutions of the sample were prepared; 10^{-1}, $10^{-1.5}$, 10^{-2}, 10^{-3}, 10^{-4} and 10^{-5}. From each dilution 200 µl was plated on Luria Bertani medium (LB) plates [1%(w/v) tryptone, 0.5% (w/v) yeast extract, 0.5% (w/v) NaCl, 1.5% (w/v) agar, pH7.5) supplemented with 1% (w/v) glucose] and incubated for four days at 37°C (Merck, Germany). Morphologically different bacteria were selected for screening and construction of a multi-genomic library.

2.2 Construction of a multi-genomic library

Morphologically different isolates were identified on Luria Bertani agar media (LB-Agar). The isolates were grown at 37°C to saturation and cultures were centrifuged at 4000 x g at 4°C for 10 min. Genomic DNA (gDNA) was extracted using the Zippy soil/fungal gDNA extraction kit according to manufacturer's instructions (ZYMO Research, USA). The small insert genomic library used for screening was constructed by Mr. Kyle Willard.

Extracted gDNA from the cultures were pooled and used to construct the library. A total of 8 µg gDNA was partially digested with 0.5 U of *Bam*HI (Thermo Scientific, USA) for 30 min yielding DNA fragments between 3 and 7 kb. DNA fragments were purified using the NucleoSpin® Extract II Kit as per manufacturer's recommendations (Clonetech Laboratories, Takara Bio, Canada). The library was created using the Stratagene Lambda Zap Express® kit (Agilent Technologies, USA), according to manufacturer's recommendations and protocol. In brief, fragments were ligated into the λ Zap II vector using the T4 DNA Ligase (Agilent Technologies, USA) and then packaged into phages. Phages of the library (10^8 Pfu/µl) was mixed with 10^9 Pfu/µl helper phages and incubated with 1 ml (10^8) XL1-Blue *Escherichia coli* (*E. coli*) cells (Agilent Technologies, USA) for

30 min at 37°C. Cell cultures were diluted in 10 ml of LB medium supplemented with Kanamycin (50 µg/ml) and incubated for 3 hrs at 37°C. The phagemid-*E. coli* cultures were then heated at 70°C for 20 min and alloquoted into 2 ml microfuge tubes. Cell lysate was centrifuged at 4000 x g for 10 min to eliminate cell debris. Phagemids were incubated with 1 ml (10^8) XL1-Blue *E. coli* cells for 30 min at 37°C. This was then plated onto 15 x 15 cm library plates with LB-Agar plates and incubated overnight at 37°C. The colonies were washed off the plates and plasmids were extracted using the alkaline lysis method (Sambrook *et al.,* 1989).

2.3 Functional screening of the genomic library for general lypolitic and acetyl esterase

Screening of the library was done on LB-Agar plates containing esterified lipids. The library was transformed into chemically competent *E. coli* DH5α (Sambrook *et al.,* 1989) and initially screened for general lypolitic (esterase/lipase) activity on tributyrin containing LB-Agar plates [1%(w/v) Tryptone, 0.5% (w/v) yeast extract, 0.5% (w/v) NaCl, 1%(w/v) Gum Arabic, 1% (v/v) Tributyrin in 1.5% Agar, pH7.5] (Kouker and Jeager 1987). Transformed cells were selected on 50 µg/ml Kanamycin and 0.1mM IPTG at 30°C for 4 days. Activity was determined by clearing zones surrounding the colonies. Plasmid DNA was obtained from positive clones using the Genejet Plasmid Miniprep kit (Thermo Scientific, USA) according to manufacturer's instructions. The extracted plasmid DNA was then transformed via heat shock into chemically competent *E. coli* DH5α (Sambrook *et al.,* 1989) and plated on Tributyrin containing LB-Agar plates and incubated at 30°C for 4 days to confirm general lypolitic activity. Clones that showed activity were further screened for acetyl esterase specific activity.

Positive clones were pooled to efficiently screen large numbers of clones for acetyl esterase activity. Clones were first inoculated, five per vial into 5 ml of LB medium supplemented with kanamycin (50 µg/ml) and grown overnight at 37°C. Pooled clones were centrifuged at 10000 x g at room temperature for 2 min to pellet the cells. The supernatant was then discarded and the pellet resuspended in 500 µl potassium phosphate buffer (50 mM, pH 7.2). The suspension was sonicated for 3 x cycle burst at 50% amplitude for 30sec to disrupt cells. Cell debris was collected by centrifugation at 10000 xg at 4°C for 20 min, and the supernatant was transferred to a sterile microcentrifuge tube.

Acetyl esterase activity was determined by incubating 100 µg of total protein extract with 1mM p-nitrophenyl acetate (pnp-OAc) (Sigma-Aldrich, Germany) in 50 mM Phosphate buffer (pH 7.2) for 10 min at 37°C. Crude protein extracted from untransformed *E. coli* was used as a negative control. The release of p-nitrophenol was measured spectrophotometrically after 10 min of incubation at a wavelength of 490nm. Thereafter pooled clones that showed acetyl esterase activity were individually cultured and individually screened for activity as described above. Plasmid DNA was extracted from the positive clones and inserts sequenced using the T7 forward primer (5'TAATACGACTCACTATAGGG3') and the M13 reverse primer (5'AACAGCTATGACCATG3') (Central Analytical Facility, Stellenbosch University, South Africa) (Oetting *et al.,* 1995).

2.4 Functional screening of isolates used to generate the genomic library

A separate set of experiments showing which of the isolated bacteria used to construct the genomic library had acetyl esterase activity they were subjected to a lypolitic activity screen as described previously with slight modification (section 2.3). Isolates were screened in the absence of IPTG and incubation time was shortened to two days.

Isolated bacteria were grown at 30°C on LB-Agar containing 1% beech wood xylan for 48 hours. Cell cultures were collected and placed into sterile 2 ml microcentrifuge tube and washed with an equal volume of 100 mM potassium phosphate buffer (pH 7.2). The washed cell cultures where centrifuged at 3000 x g for 10 min at 4°C to collect the cells. The cells were resuspended in an equal volume of 50 mM potassium phosphate buffer (pH 7.2) and the suspension was sonicated for 3 x cycle burst at 50% amplitude for 30 sec. Cell debris was removed by centrifugation at 10000 x g at 4°C for 20 min. The supernatant was inserted into a sterile microcentrifuge tube and protein concentration determined by Bradford assay (Bradford et al., 1976). This crude protein fraction was assayed for acetyl esterase activity by incubating 50 µl total protein extract with 1 mM p-nitrophenyl acetate (pnp-OAc) (Sigma-Aldrich, Germany) in 50mM phosphate buffer (pH 7.2) for 10 min at 37°C. Crude protein extracted from untransformed *E. coli* was used as a negative control. The release of p-nitrophenol was measured after 10 min spectrophotometrically at 490nm.

2.5 Sequence analysis and annotation

Sequences obtained from clones showing acetyl esterase activity were assembled into contigs using BioEdit applying the Clustel W algorithm (Thompson *et al.,* 1994). Open reading frames were identified using Softberry (http://linux1.softberry.com/berry.phtml), ORF Finder (http://www.ncbi.nlm.nih.gov/projects/gorf) and EMBL-EBI database (http://www.ebi.ac.uk/). Open reading frames (ORF) coding for acetyl esterases were analysed using the non-redundant protein sequences (nr) database using NCBI Blastp algorithm (http://blast.ncbi.nlm.nih.gov) and UniprotKB (http://www.uniprot.org/) databases. Four sequences with significant similarity to acetyl hydrolases and acetyl esterases were selected for further study.

2.6 Acetyl esterase phylogenetic analysis

Amino acid sequences of the four isolated enzymes showing acetyl esterase activity were aligned with 100 of their closest relatives within GenBank and EMBL databases using standard parameters, after which they were grouped and aligned using Clustel W (Thompson *et al.,* 1994). This created clades of similarity and 3 of the sequences most similar to the query were used to constitute the branches of the final tree. These sequences where then re-aligned with subsequent groups and Neighbour-joining (N-J) Maximum likelihood (ML) trees drawn using MEGA 5 (Tamura *et al.,* 2011). Parameters and algorithms for tree construction was calculated using jModeltest (Posada, 2008), a cut-off of 95% parameter fit on Akaike Information Criterion (AIC) was implemented to define significance. Significance of trees was calculated by 1000 bootstrap replications, all bootstrap values above 50% where deemed significant and reported (Holmes, 2003). N-J and ML trees were drawn and bootstrap values for both were incorporated on one tree.

2.7 Cloning of the acetyl esterase coding regions

Primers were designed for the amplification of the open reading frames coding for the four acetyl esterase enzymes identified by sequencing and AE screening from the library clones 1, 12, 21 and 37. To amplify the genes, the PCR reaction included 5 ng of DNA, 1 X GoTaq green master mix buffer [containing 1.5 mM MgCl, 0.2 mM of dNTPs, 0.5 U Taq DNA polymerase (Promega, Germany)], 0.5 µM forward and 0.5 µM reverse primers (Table 2A).

The PCR parameters were as follows; initial denaturation temperature of 95°C for 2 min, 95°C for 30 sec, annealing temperature of 58°C for 45 sec and an elongation step of 72°C for 45 sec repeated for 25 cycles prior to a final extension step of 5 min at 72°C and cooled at 4°C for 5 min. The PCR product was subjected to electrophoresis on a 0.8% agarose gel and purified using the Nucleospin II gel extraction kit as per manufacturer's instructions (Clonetech Laboratories, Takara Bio, Canada). Amplicons from the successful PCR were designated *HEL1*, *HEL12*, *HEL21*, and *HEL37* (Table 2A). Fragments were then ligated into pGem-T-easy, according to manufacturer's specifications (Promega, Germany). Transformation reaction was performed by transforming 5 µl of the ligation reaction into chemically competent *E. coli* DH5α (Sambrook *et al.,* 1989), and plated onto LB containing 100 µg/ml Ampicillin. The plates were incubated overnight at 37°C. Insert size was confirmed by screening colony PCR. Amplified PCR product was subjected to electrophoresis on a 0.8% agarose gel and positive clones where inoculated overnight in 2 ml LB with 100 µg/ml Ampicillin for selection. The overnight culture was then used for a plasmid extraction with a plasmid miniprep kit (Thermo-Scientific).

DNA fragments were excised from pGEM-T-easy:*HEL12* and pGEM-T-easy:*HEL37* constructs with 0.1 U of *Bgl*II (Thermo Scientific, USA) and 0.2 U of *Kpn*I (Thermo Scientific, USA) according to manufacturer's instruction. The gel excised fragments were purified using Nucleospin II gel extraction kit, according to manufacturer's instructions (Clonetech Laboratories, Takara Bio, Canada) and ligated into bacterial expression vector pRSET-C (Invitrogen, USA) digested with 0.1 U of *Bgl*II (Thermo Scientific, USA) and 0.2 U of *Kpn*I (Thermo Scientific, USA). Ligation reactions were conducted at 4°C overnight and transformed into chemically competent *E. coli* DH5α (Sambrook *et al.,* 1989). Transformed cells were plated on LB-Agar containing 100 µg/ml Ampicillin for selection and incubated at 37°C overnight. PCR amplification was performed on five colonies from each plate using the following conditions: 1X GoTaq green master mix buffer, 0. 5µM forward (5'TAATACGACTCACTATAGGG3') and 0.5 µM reverse primers (5'CCCCGTCGAC CTAAAGGTGTGTGGTAAAAA3' and 5'CCCCCGTCGACGGTTAGTAGAAGAATAGG3') respectively. Amplified PCR product was subjected to electrophoresis on a 0.8% agarose gel. Positive clones were inoculated in 2 ml LB containing 100 µg/ml Ampicillin for selection and incubated at 37°C overnight. Plasmid extraction of the inoculated colonies was done with Thermo Scientific plasmid miniprep kit (Thermo Scientific, USA). Ten nanograms of the plasmid was sent for sequencing using the T7 reverse (5'GCTAGTTATTGCTCAGCGG3')

and T7 Promoter (5'TAATACGACTCACTATAGGG3') sequencing primers to confirm the reading frame and sequence (Central Analytical Facility, Stellenbosch).

2.8 Protein purification and enzyme characterisation

2.8.1 Protein purification

The pRSET C:6xHis::*HEL12* and pRSET C:6xHis::*HEL37* constructs were transformed via heat shock into *E.coli* BL21 (DE3) PlysS (Promega, Germany) (Kleber-Janke and Becker 2000; Sambrook *et al.,* 1989). Transformations were plated onto LB-Agar (Merck), containing 100 µg/ml Ampicillin and 35 µg/ml Chloramphenicol. The transformation of both constructs into the *E.coli* BL21 (DE3) PlysS (Promega, Germany) was confirmed by PCR. Positive colonies containing each transformed construct was then inoculated into a 10 ml LB (Merck) containing 100 µg/ml Ampicillin and incubated for 5 hours to an OD_{600} = 0.100 at 37°C. Cell cultures were centrifuged at 2000 x g for 20 min at 4°C, the supernatant discarded and the pellet resuspended in LB containing 100 mg/ml Ampicillin and grown to an OD_{600} = 1.000. A 2 ml aliquot was transferred aseptically into a sterile 2ml microfuge tube for an acetyl esterase screen. Aliquots were centrifuged at 10000 x g at room temperature for 2 min to pellet the cells. Crude protein extracted from untransformed *E. coli* was used as a negative control. The release of p-nitrophenol was measured after 10 min spectrophotometrically at 490 nm. Once AE activity was confirmed cell cultures were centrifuged at 2400 rpm for 20 min at 4°C. Two 500 ml LB (Merck) culture media with 100 µg/ml Ampicillin and 0.1M IPTG in 1 L shaker flasks was inoculated with each of the positive clones to OD_{600} = 0.100. This was cultured overnight at 28°C without IPTG. Overnight cultures were centrifuged at 3500 x g for 20 min at 4°C to harvest cells. The supernatant was discarded and the pellet resuspended in 10 ml of 100 mM potassium phosphate buffer (pH6.8) containing 40 mM imidazole.

The suspension was subjected to 3 x sonication cycles at 50% amplitude on ice for 30 sec and placed on ice for 1 min. Sonicated cells were centrifuged at 8000 x g for 25 min to remove cell debris. Supernatant was filtered through a 0.45µm filter and then also filtered using a 0.22 µm filter to remove small cell wall particulates. The sample was kept cold throughout the protein preparation process.

Recombinant protein was purified using a HistrapTM HP 5ml column (GE Healthcare Life Sciences, USA) using an ÄKTA FPLC purifier using potassium phosphate pH 6.8 (GE Healthcare Life Sciences, USA) as running buffer. Elution of the bound recombinant protein was achieved with a gradient from 0% - 100% (100 mM potassium phosphate at pH 6.8 with 40 mM Imidazole) to B (100 mM potassium phosphate at pH 6.8 with 400 mM Imidazole). Purification of proteins were visualised by 10% SDS-Page electrophoresis and protein concentration determined by Bradford assay (Bradford et al., 1976). Fractions that contained the purified protein was dialysed using 5.0K MWCO cellulose membranes against ddH2O overnight at 4°C to remove the remaining Imidazole. Dialysis was repeated against 100 mM potassium phosphate overnight at 4°C to facilitate buffer exchange. Glycerol was added to the protein to a concentration of 10% and stored at 4°C.

2.8.2 Standard Enzyme Activity Assay

The ability of acetyl esterase (AE) enzymes to cleave ester bonds was measured using the synthetic compound p-nitrophenyl acetate (Biely *et al.,* 1991). Release of p-nitrophenol by cleavage of the ester bond was continuously monitored at 490 nm. The substrate p-nitrophenyl acetate was dissolved in 100% ethanol to a concentration of 100 mM. Substrate was then diluted to 40 mM with 100 mM pH 7.0 potassium phosphate and used immediately. Standard acetyl esterase assay (AE assay), unless otherwise stated was conducted in 190 µl 50 mM potassium phosphate buffer (pH 7.0) (Merck, Germany), with a final concentration of 2 mM p-nitrophenyl acetate. To initiate the reaction 10 µg (10 µg/µl) purified protein extract was added (Levisson *et al.,* 2012). Reaction was continuously monitored at OD_{490} for 10 min at a temperature of 30°C, taking measurements every 15 sec. An assay is defined as six technical replicates repeated with two in biological repeats. The assay was repeated without enzyme to construct a degradation baseline to normalise assay values. p-Nitrophenol was used to set up a standard curve of absolute hydrolysis.

2.8.3 Substrate stability

In order to establish the degree of instability of p-nitrophenyl acetate and recognise possible solutions, stability assays were conducted on p-nitrophenyl acetate in different buffers and solutes. Each data point presented in experimental results represents ±S.E.M of 8 independent replicates. The assay was repeated without enzyme to construct a degradation baseline to normalise assay values.

Stability Assay 1

Evaluation of the degradation of p-nitrophenyl acetate in four solutions of the substrate and control substrate (p-nitrophenol) was prepared to saturation using solutes: 100% ethanol, 20% ethanol, DMSO (Dimethylsulfoxide) and DMF (Dimethylformamide) at 10 mM and 100 mM concentrations. The substrates where then diluted to 2 mM concentrations in Potassium phosphate buffer, pH6.8 in a volume of 210 µl and subjected to spectrophotometric assays at 490nm (Levisson *et al.,* 2012). Measurements were taken at 0, 10 min, 20 min, 60 min, 24 hrs, 48 hrs, 72 hrs and 96 hrs. Each data point presented in experimental results represents ±S.E.M of 8 independent replicates. The assay was repeated without enzyme to construct a degradation baseline to normalise assay values.

Stability Assay 2

The effect of the change in pH on the p-nitrophenyl acetate was determined by using the same protocol as described in Stability Assay 1 with slight modification. Two buffer systems used, potassium phosphate and sodium phosphate at a pH of was 4.7, 7.2, and 9.8. Assay was assessed at an endpoint of 10 min. Each data point presented in experimental results represents ±S.E.M of 8 independent replicates. The assay was repeated without enzyme to construct a degradation baseline to normalise assay values.

2.8.4 Substrate specificity

The specificity for the maximum carbon chain length was determined by using five different substrates; p-nitrophenyl acetate (3 carbon), p-nitrophenyl butyrate (5 carbon),

p-nitrophenyl octanoate (8 carbon), p-nitrophenyl dodecanoate (12 carbon), and p-nitrophenyl palmitate (16 carbon) in the standard acetyl esterase screen as described in Section 2.8.2. Ethanol is not a solvate that can dissolve p-nitrophenyl dodecanoate and p-nitrophenyl palmitate, and they were therefore dissolved in DMF and used immediately. Each data point from each assay represents a mean ±S.E. of 8 technical repeats by two biological repeats. The data from each assay was normalised by a baseline determined by running said assay to completion without adding enzyme to measure the extinction constant of each substrate respectively.

2.8.5 Optimum pH for HEL12 and HEL37

Determination of optimum pH was done by measuring p-nitrophenol release in 50 mM buffer systems ranging from pH 3.0 to pH 10.0. phosphate-citrate buffer (pH 3.0-6.5), potassium phosphate buffer (pH 6.0-8.0), glycine-KOH buffer (pH 8.0-10.0) was used. Activity was observed as a continuous assay using standard acetyl esterase protocol described in section 2.8.2. The data points presented represent a mean ± S.E. of 8 repeats and a two biological repeats. The data points was normalised by subtracting baseline values collected by repeating the assay without enzyme.

2.8.6 Optimum Temperature for HEL12 and HEL37

Acetyl esterase activity was established at a temperature range between 25°C to 70°C. Assays were conducted in 50 mM potassium phosphate (pH 7.0). The temperature optima assay was conducted with 1140 µl 50mM potassium phosphate buffers (pH 7.0), 2 mM p-nitrophenyl acetate and 54 µg of enzyme (10 µg/µl). Firstly, assay buffer was heated for 10 min to correct temperature before p-nitrophenyl acetate was added. After which the assay was initiated by adding enzyme. Temperature assay was respectively conducted for 10 min at each of the different temperatures, and terminated by placing the reaction in an ice-ethanol bath. Absorbance values of completed assays was subsequently analysed in a total volume of 200 µl at 490 nm. The data points presented represent a mean ± S.E. of 8 repeats and two biological repeats. The data points was normalised by subtracting baseline values collected by repeating the assay without enzyme.

2.8.7 Kinetic assay for HEL12 and HEL37 Acetyl esterase activity

The kinetic parameters of the acetyl esterase enzymes was assessed with increased concentrations of p-nitrophenyl acetate from 0 mM to 14 mM using the reaction conditions established for the standard acetyl esterase activity assays described in section 2.8.2. HEL12 and HEL37 kinetic data were processed with Prism 5 (Graphpad, USA). K_m and V_{max} values were calculated by using non-linear Michaelis-Menten regression plots in conjunction with Lineweaver-burk plots. Prism 5.0 (Graphpad, USA) was used to calculate the turnover number k_{cat} = V_{max}/Molar concentration of HEL12 and HEL37. Data points were normalised using a water control. The data points presented represent a mean ± S.E. of 8 repeats and two biological repeats.

2.8.8 Native Substrate Assay

Native substrate for acetyl xylan esterases is acetylated xylan polymers. The investigation of polymer substrate affinity an assay was designed with 50 mg/ml stock solutions of the following native substrates: (1) wheat arabinoxylan (Megazyme), (2) xylan from oats (Sigma-Aldrich), (3) arabinan from sugar beet (Megazyme), (4) xyloglucan from tamarind seed (Megazyme), (5) sucrose acetate isobutyrate (Sigma-Aldrich), (6) chitin from shrimp (Sigma-Aldrich), (7) xylan from Beechwood (Sigma-Aldrich) in 1% DMSO solution and 50 mM potassium phosphate buffer (pH7.4) (Merck, Germany). Assay was conducted in 190 µl 50 mM potassium phosphate buffer (pH7.4) (Merck, Germany), with a final concentration of 2.5 mg/ml of substrate. To initiate the reaction 10 µg (10 µg/µl) purified protein extract was added (Levisson *et al.,* 2012). Reaction was incubated at 30°C for 10 min. Hydrolysis reaction was terminated by cooling sample to 0°C. Acetic acid was quantified by use of acetic acid assay as per microplate assay protocol (Megazyme, K-ACETRM). The assay controls were conducted (1) without enzyme to construct a degradation baseline to normalise assay values, (2) with boiled enzyme as negative control (3) each substrate was assayed as above, with crude protein extract from untransformed *E. coli* BL21 and (4) a final control without substrate to indicate optical density changes due to addition of enzyme. The data points presented represent a mean ± S.E. of 8 repeats and two biological repeats.

2.9 Identification of the bacterial strains used for library construction

Bacterial isolates C1, C3, C7, C8, C10, K2, K3, K4, K6, K8 were incubated in LB supplemented with glucose (Kirk *et al.,* 2004). The cultures were streaked onto LB-Agar plates supplemented with 0.1% beechwood xylan to induce expression of acetyl xylan esterase genes. Streak plates were incubated at 37°C overnight. Cultures were scraped from the plates to perform a crude screen for acetyl esterase activity. Cells were spotted onto a square of nitro-cellulose blotting paper saturated with a solution of 2mM p-nitrophenyl acetate (Sigma) in 100 mM potassium phosphate (pH 6.8). The back of the nitro-cellulose blotting paper was visualised for yellow halo formation indicating the presence of acetyl esterase enzymes.

Genomic DNA (gDNA) was extracted from the isolates using the Zippy soil/fungal gDNA extraction kit (Zymo Research, USA). The extracted gDNA was used for 16S PCR amplification, 5 ng of gDNA per reaction was used to amplify the 16S ribosomal DNA (rDNA). Primers E9F, U529 (primer set 9 in Table 2A) F8 and 1512 (Primer set 10 in Table 2A) amplified 560 bp and 1487 bp fragment of the 16S rDNA respectively (Baker *et al.,* 2003). PCR reaction included 5 ng of gDNA, 1x Mytaq red reaction buffer 0.5 U Mytaq polymerase, 0.5 µM forward and 0.5 µM reverse primers. Volume of the reaction was adjusted to 25 µl. The PCR parameters were as follows: initial denaturation of 15 min at 95°C, 25 cycles of 1 min at 95°C, 45 sec at 55°C annealing temperature, and 1 min at 72°C. With a final extension of 72°C for 15 min, the reaction was cooled down to 4°C.

The PCR product was subjected to electrophoresis on a 0.8% agarose gel and purified from the gel using Nucleospin II gel extraction kit as per manufacturer's instructions (Clonetech Laboratories, Takara Bio, Canada) and ligated into pJET (Thermo Scientific, USA) according to manufacturer's instructions. These constructs were then transformed via heat-shock into chemically competent *E. coli* DH5α cells (Sambrook *et al.,* 1989). Single colonies were picked up and inoculated into 2 ml of LB with 100 mg/l Ampicillin and incubated at 37°C overnight. Plasmid DNA from the 2 ml overnight cultures was prepared using the Genejet Plasmid Miniprep kit (Thermo Scientific, USA). DNA sequencing was performed by Central Analytical Facility, Stellenbosch (Stellenbosch University, Stellenbosch, South Africa).

NCBI BLAST (blastn, http://blast.ncbi.nlm.nih.gov) was used to analyse the 16S rDNA sequences originating from the isolates (Altschul *et al.,* 1990). After the initial sequence

profiling the sequences were grouped and aligned with Clustel W (Thompson *et al.,* 1994) using BioEdit (Hall, 1999). The top ten BLAST hits were used to cluster the isolates to evaluate their phylogenetic origins. These were then re-aligned with subsequent groups and Neighbour-Joining (NJ) trees drawn using MEGA 5 (Tamura *et al.,* 2011) upon which classification was inferred (Saitou and Nei, 1987). Parameters and algorithms for tree construction were calculated using jModeltest (Posada, 2008). A cut-off of 95% parameter fit on Akaike Information Criterion (AIC) was implemented to define significance. MEGA was used to draw phylogenetic trees (Tamura *et al.,* 2011). Significance of trees was calculated by 1000 bootstrap replications. All bootstrap values above 50% were deemed significant and reported (Holmes, 2003). N-J and ML trees were drawn and bootstrap values for both were incorporated on one tree.

Chapter 3

3. Results

Prior to creating a multigenomic library, isolated bacteria were screened for lypolitic and acetyl esterase (AE) activity. This facilitated defining the isolates that were incorporated into this multigenomic library and screening the isolated bacteria for general lypolitic activity.

3.1 Screening isolates used for library construction for esterase/lipase activity

Bacterial isolates used to create the small insert library were screened for lypolitic and esterase activity (Table 3.1). The screen resulted in eight morphologically different isolates based on their ability to degrade simple lipids and p-nitrophenyl acetate. Seven of these isolates (C7, C8, C10, K2, K3, K4, K6, K10) showed acetyl esterase (AE) activity (Table 3.1) when using p-nitrophenyl acetate as substrate. In addition to showing acetyl esterase activity seven (C1, C3, C7, C8, K4, K6, K10) also showed tributyryl esterase activity. Three strains, designated C10, K2, and K3 showed no esterase/lipase activity, but were positive for acetyl esterase activity. Five isolates (C7, C8, K4, K6, K10) showed both AE and lipase/esterase (tributyl esterase) activity.

Table 3.1. Summary of isolated bacteria with designations and their individual esterase/lipase and AE activity. (-) indicates no activity, (+) indicates positive activity and (++) designates high activity]

Name of isolate	Tributyryl Esterase/lipase Activity	Acetyl esterase activity (AE)
C1	++	-
C3	+	-
C7	+	+
C8	+	+
C10	-	++
K2	-	+
K3	-	+
K4	+	+
K6	+	+
K10	+	+

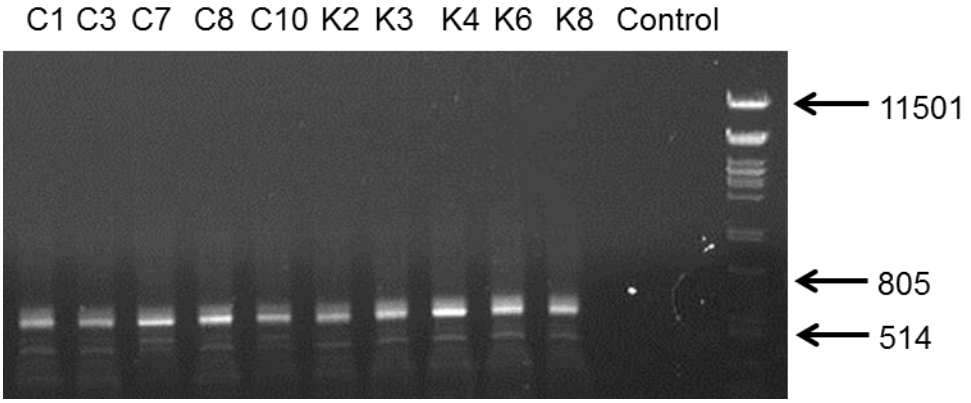

Figure 3.1. 16S rDNA amplification of bacterial isolates for phylogenetic analysis. PCR amplification was performed on genomic DNA extracted from isolates C1, C3 C7, C8, C10, K2, K3, K4, K6 and K8, respectively. The individual 16S rDNA was used to construct a phylogenetic tree.

Isolated bacteria were genotyped using 16S PCR analysis (Table 3.3). Closets relatives were *Acinetobacter sp. LC084* (C8), *Bacillus sp.* (C10, K6, K8), *Klebsiella pneumoniea* (K2), *Chronobacter sakazakii* (K3) and *Citrobacter farmeri* (K4). Lipolytic activity has been described in C3, *Acinetobacter sp. LC084* (C8), *Bacillus sp.* (C10, K6, K8), *Klebsiella pneumoniea* (K2) and *Chronobacter sakazakii* (K3). Esterase activity has been documented in *Staphyococcus sp. IARI-S-21* (C1), *Bacillus pumillis* (C7), *Acenetobacter sp. LC084* (C8), *Bacillus sp.* (C10, K6, K8), *Klebsiella pneumoniea* (K2) and *Chronobacter sakazakii* (K3).

Table 3.2: BLAST analysis of 16S rDNA sequences of the isolated bacteria. Nearest relatives to bacterial isolates were tabulated.

Co	Cellular Lipase
C1	No activity reported
C3	No activity reported
C7	No activity reported
C8	Shown to have activity (Mccoy 2011)
C10	Most strains have activity (Han SI 1995)
K2	Most strains have activity (Han SI 1995)
K3	Most strains have activity (Yu EK, 1985)
K4	Most strains have activity (Grimm M, 2008)
K6	Shown to have activity (Mccoy 2011)
K8	Shown to have activity (Mccoy 2011)

Co	Cellular Lipase
C1	No activity in literature
C3	No activity in literature for species, only genus (Thermostable Lipase reported by Sharif S 2011)
C7	No activity in literature
C8	Shown to have activity (Mccoy 2011)
C10	Shown to have activity (Lakhanpal RM, 2011)
K2	Shown to have activity (Lakhanpal RM, 2011)
K3	Shown activity in different strain (Sekowska D, 2006)
K4	no activity in literature
K6	Mccoy 2011)
K8	Accession 2011)

Co	Org	Lipase Ester	Cs	ae
C1	*Staphylococcus sp. IARI-S-21*	esterase (Mccoy 2011)	99%	96%
C3	*Staphylococcus epidermidis strain SR1*	no activity in literature for species, only genus	99%	96%
C7	*Bacillus pumilus strain S811*	esterase (Mccoy 2011)	93%	91%
C8	*Acinetobacter sp. LC084*	esterase (Mccoy 2011)	97%	95%
C10	*Bacillus sp.*	esterase (Mccoy 2011)	98%	95%
K2	*Klebsiella pneumoniae strain BR10_2A*	esterase activity in genus (Wang and Tsai, 2009)	93%	98%
K3	*Cronobacter sakazakii*	esterase (Mccoy 2011)	98%	94%
K4	*Citrobcter farmeri*	no activity in literature	98%	89%
K6	*Bacillus subtilis strain NA12*	esterase (Mccoy 2011)	99%	92%
K8	*Bacillus selenatarsenatis*	esterase activity in genus (Perrson and Bornscheur 2003)	97%	96%

Table 3.3: Isolates were classified using standard microbial techniques. Results show the class of bacterium, and indicate the gram stain, catalase activity, oxidase activity, oxygen tolerance.

Class	Gram stain	Catalase activity	Oxidase	Oxygen Tolerace	Form	Designation	Reference
Kingdom: *Bacteria* Phylum: *Firmicutes* Class: *Bacilli* Order: *Bacillales* Family: *Staphylococcaceae* Genus: *Staphylococcus*	Positive	Positive	Negative	facultative anaerobe	Round Cocci	C1, C3	Takahashi *et al.,* 1999
Kingdom: *Bacteria* Phylum: *Firmicutes* Class: *Bacilli* Order: *Bacillales* Family: *Bacillaceae* Genus: *Bacillus*	Positive	Positive	Negative	aerobes or facultative anaerobes	Rod-shaped	C7, C10, K6, K8	Alcaraz *et al.,* 2010
Kingdom: *Bacteria* Phylum: *Proteobacteria* Class: *Gammaproteobacteria* Order: *Pseudomonadales* Family: *Moraxellaceae* Genus: *Acinetobacter*	Negative	Positive	Negative	aerobe	Round	C8	Visca *et al.,* 2011
Kingdom: *Bacteria* Phylum: *Proteobacteria* Class: Gamma *Proteobacteria* Order: *Enterobacteriales* Family: *Enterobacteriaceae* Genus: *Klebsiella*	Negative	Positive	Negative	facultative anaerobe	Rod-shaped	K2	Podschun and Ullmann, 1998
Kingdom: *Bacteria* Phylum: *Proteobacteria* Class: *Gamma Proteobacteria* Order: *Enterobacteriales* Family: *Enterobacteriaceae* Genus: *Cronobacter*	Negative	Positive	Negative	facultative anaerobe	Rod-shaped	K3, K4	Iverson *et al.,* 2008

Phylogenetic analysis shows four distinct branches to the N-J/M-L phylogenetic tree depicting the ten bacterial isolates (Figure 3.2). A 600bp fragment of the VII region of the 16S ribosome was amplified and sequenced and these sequences were related to their closest ATCC curated reference sequences. Four isolates were found to be related to *Bacillus* strains: (C7) *B. punius,* (K6) *B. vallismortis,* (K8) *B. jeotgali* and (C10) *B. thuringiensis*. The other Isolates that were part of the library, branched off on a different branch and were closely related to (C1) *Staphylococcus pastori* and (C3) *S. epidermis*. The last four isolates clustered into a completely different clad, with (C8) *Acinetobacter lwoffii*, (K3) *Serraria nematodiphila*, (K2) *Enterobacter kobei* and a connation of (K4) being related to *Citrobacter farmer*.

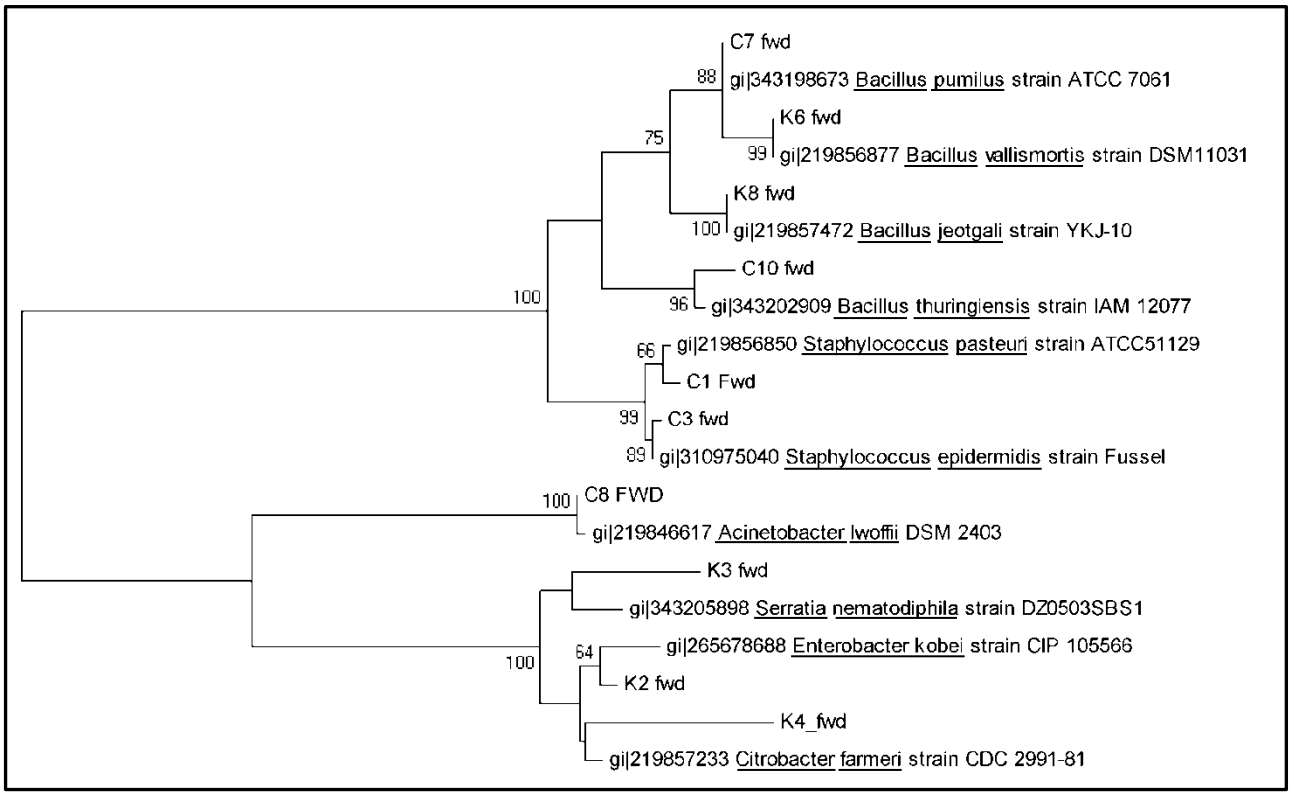

Figure 3.2. DNA sequence of 16S rDNA combining Neighbour-joining (N-J) and Maximum Likelihood (M-L) phylogenetic trees alignment indicating similarity to known ATCC curated sequences from known bacterial isolates.

3.2 Screening small insert plasmid library for esterase/lipase and AE enzyme activity using *E. coli* DH-5α

Plated assay screen was done on 25 x 25cm square screening plates, on each plate 10 000 clones were screened. With a total of 300 plates screened, 3 million clones were screened for esterase/lipase activity on tributyrin containing media. Clones showing general lypolitic activity were identified by a clearing zone of around each colony. Plate assays identified 53 positive clones. The clearing zones for each clone varied in size (Figure 3.3). Clones that produced the largest zones of clearance were chosen for further study.

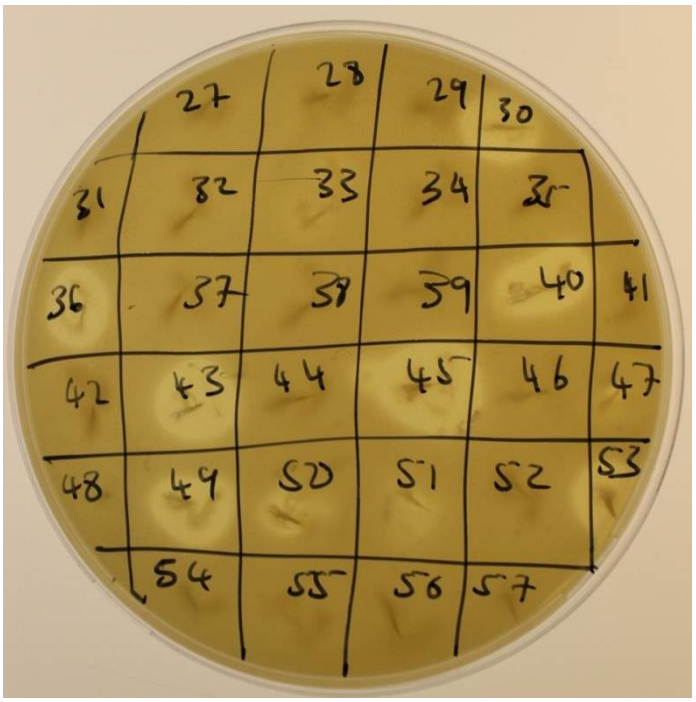

Figure 3.3. Tributyrin plate assay for general lypolitic activity. Activity is indicated by zones of clearing surrounding positive clones. Clones 30, 36, 40, 43, 45, 49, 50, 51 and 53, show general lypolitic activity.

3.2.1 Identification of clones with Acetyl Esterase activity

In order to establish greatest activity amongst the isolated clones the 53 clones were grouped into 11 groups of 5 each and one with 3 clones, and screened for acetyl esterase activity. Groups 2, 6, 9, 10, and 12 (Figure 3.4) showed the highest activity and were selected for further study (Figure 3.5).

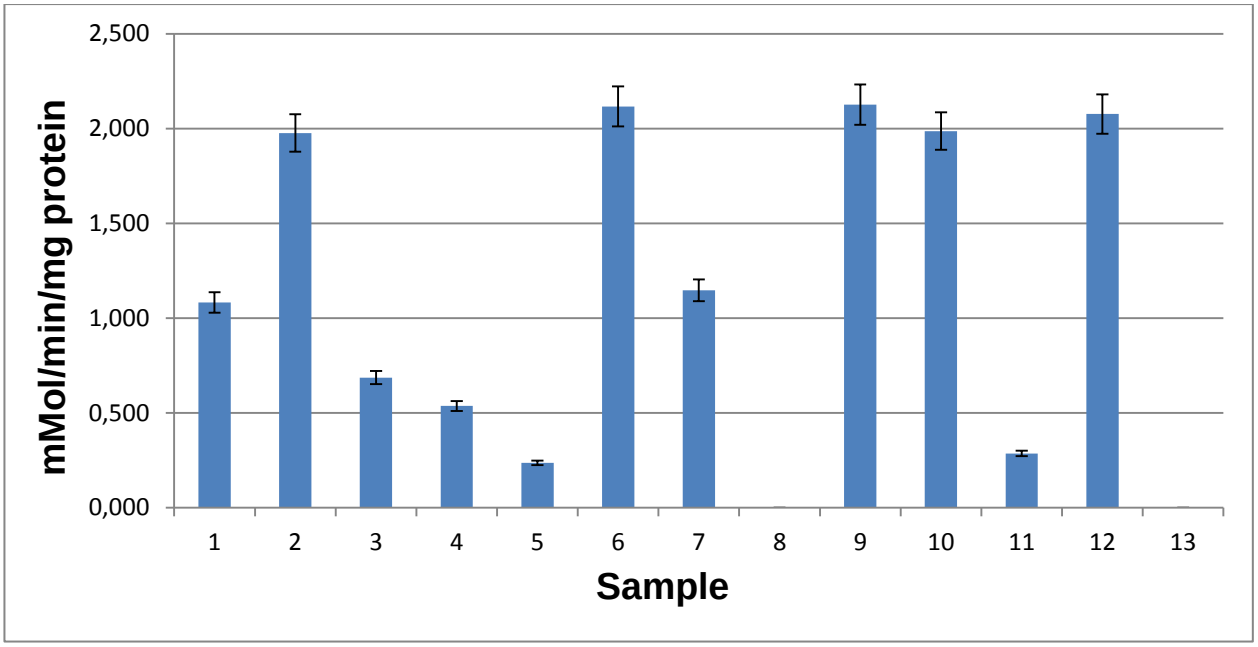

Figure 3.4. Twelve samples that where screened for AE activity on p-nitrophenol acetate. Groups 2, 6, 9, 10 and 12 screened positive for AE activity. Sample 13 was a negative control. Sample 8 was a clone that had no AE activity.

Individual clones within the groups that showed highest activity were identified and tested individually for AE activity (Figure 3.5). Clones with the most significant activities (clones 29, 37, 27, 35, 21, 1, 12.1 and 3-5) were chosen for further analysis.

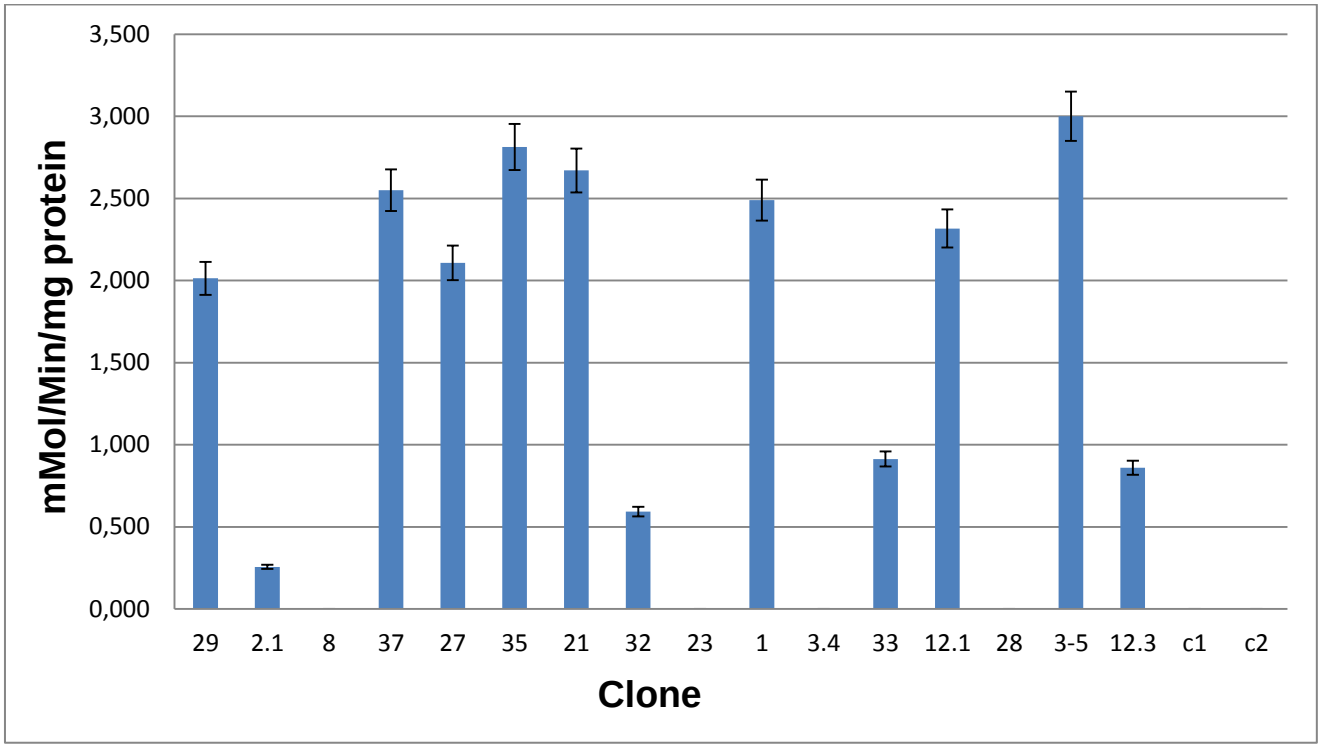

Figure 3.5. Acetyl esterase activity determination activity on p-nitrophenol acetate for individually isolated clones. Sixteen clones were screened showing varied AE activity. Activity above 2 mMol/min/mg was seen as significantly high and considered for further examination.

Plasmid preparations of clones 29, 37, 27, 35, 21, 1, 12.1 and 3-5 were restriction digested and the respective restriction patterns indicated some similarity amongst the positive clones (Figure 3.6). Based on the restriction patterns four clones (37, 1, 12 and 21) were identified as unique and selected for further study. These clones were re-designated HEL1, HEL12, HEL21 and HEL37 to indicate their origin from the hot wood library, of esterase and lipase clones. Clones HEL1, HEL12, HEL21 and HEL37 were then sequenced (CAF, Stellenbosch University, South Africa).

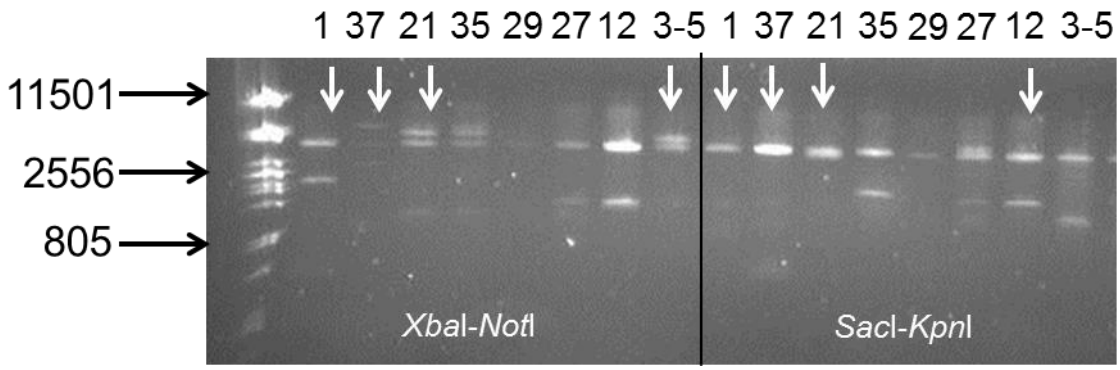

Figure 3.6. Restriction digest pattern of plasmids isolated from positive AE clones (Marker ladder: λ-*Pst*I). Lanes 1: λ-*Pst*I, Lanes 2-9: plasmids clones digested with enzymes *Xba*I-*Not*I. Lanes: 10-17 plasmid digested with *Sac*I-*Kpn*I. Four clones (indicated by arrows) were chosen for further analysis based on their unique restriction patterns HEL1 (approx. 4 kb + 2 kb, and 4.5 kb), HEL37 (uncut plasmid, and approx. 4.5 kb), HEL21 (uncut plasmid, and approx. 4.5 kb) and HEL12 (approx. 4 kb + 1.7 kb, and 4 kb + 1.7 kb) were chosen for further analysis.

Protein equences obtained from plasmid stocks of HEL1, HEL12, HEL21 and HEL37 were screened via t-BLASTX and Softberry to identify likely open reading frames (ORF) and putative genes (Table 3.4). Sequences obtained from DNA sequence analysis of HEL1, HEL37, HEL21 and HEL12 plasmids were respectively analysed using t-BLASTX and Softberry algorithms to identify likely open reading frames (ORF) and putative genes (Table 3.5). *HEL1* (862 bp) analysis returned BLAST hits to one recognised ORF, an esterase from *Enterobacter radicans* (96%). Similarly, BLAST results of *HEL37* (630 bp) also recognised one esterase ORF of approximately 789 bp in length belonging to *Staphylococcus hominis*. The *HEL21* insert (744 bp) had the highest BLAST hit to a 726bp hydrolase from *Citrobacter freundii* (72%). *HEL12* (741 bp), BLAST results returned to an esterase from *Chronobacter sakazakii* at 76% similarity.

Sequence analysis results illustrated sequence novelty for the *HEL12* and *HEL37* ORF's. On this basis it was decided to continue with further analysis of these clones as they provided the best possibility for a novel esterase activity.

Table 3.4. Gene identification based on t-BLASTX sequence analysis of library clones, showing peptide length, gene size, Nucleotide length, molecular weight (www.calctool.org), percentage (%) identity to known protein, name of protein, organism associated with protein and the E-Value linked to the protein analysis.

Designation	Peptide length (aa)	Nucleic acid length (bp)	Molecular Weight (kDa)	% Identity	Best Hit Association	Organsim Name	E-Value
HEL 1	249	747	29.88	96%	Esterase	*Entrobacter Radicitans*	2.0x10^175
HEL 12	242	744	29.04	76%	Esterase	*Cronobacter sakazakii*	2.0x10^39
HEL 21	249	725	29.88	72%	Hydrolase	*Citrobacter freundii*	3.0x10^64
HEL 37	277	831	33.24	74%	Esterase	*Staphylococcus hominis*	3.0x1-^135

Using the NCBI-BLAST algorithm, HEL37 protein (210aa) was found to be most closely related to the tributyrin esterase protein from *Staphylococcus hominis*. Alignment of this protein with the tributyrin esterase sequence from *Staphylococcus hominis* shows 72% similarity to *Staphylococcus hominis* tributyrin esterase (Table 3.4). Alignments highlighted an esterase conserved area (COG0627) at interval 1-209aa and a predicted alpha/beta hydrolase domain (COG2819) at interval 41-142 aa (http://www.ncbi.nlm.nih.gov /Structure/cdd/cdd.shtml) (Figure 3.7 and 3.8). Deletion points are present in HEL37 with regard to COG0627 (42-28 aa) and COG2819 (41-58 aa).

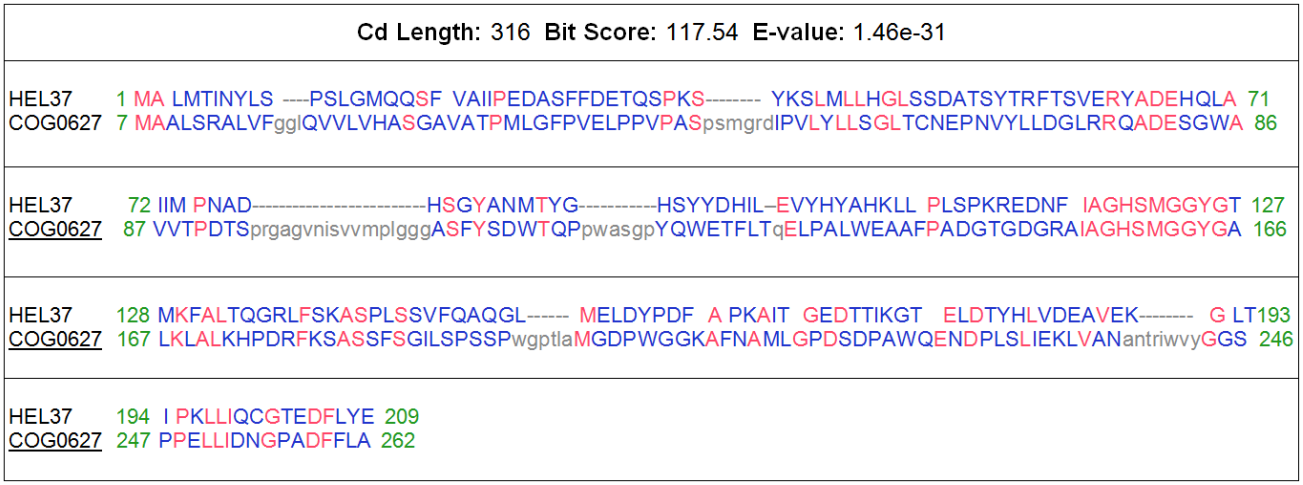

Figure 3.7. The predicted esterase conserved domain COG0627 aligned with amino acids 1-209 of HEL37. Amino acids depicted in blue refer to residues that are similar but not identical. Amino acids highlighted in red refer to conserved residues.

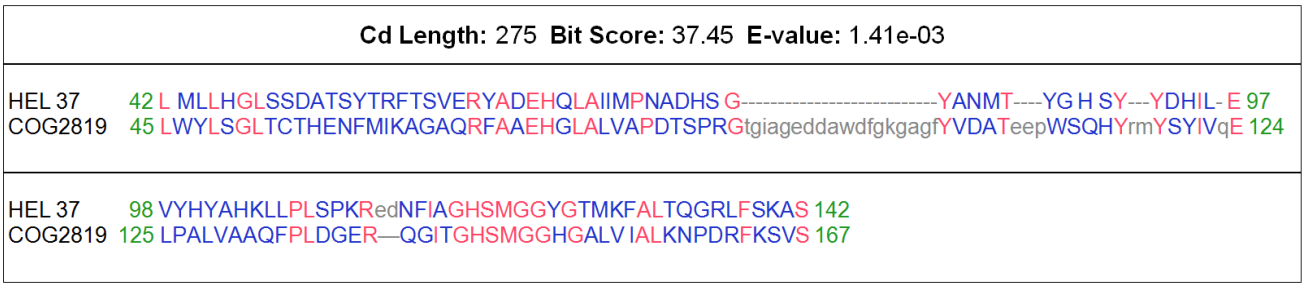

Figure 3.8. Alpha/beta hydrolase conserved domain COG2819 aligned with amino acids 42-142 of HEL37. Amino acids depicted in blue refer to amino acids that are similar but not identical. Amino acids depicted in red refer to conserved residues.

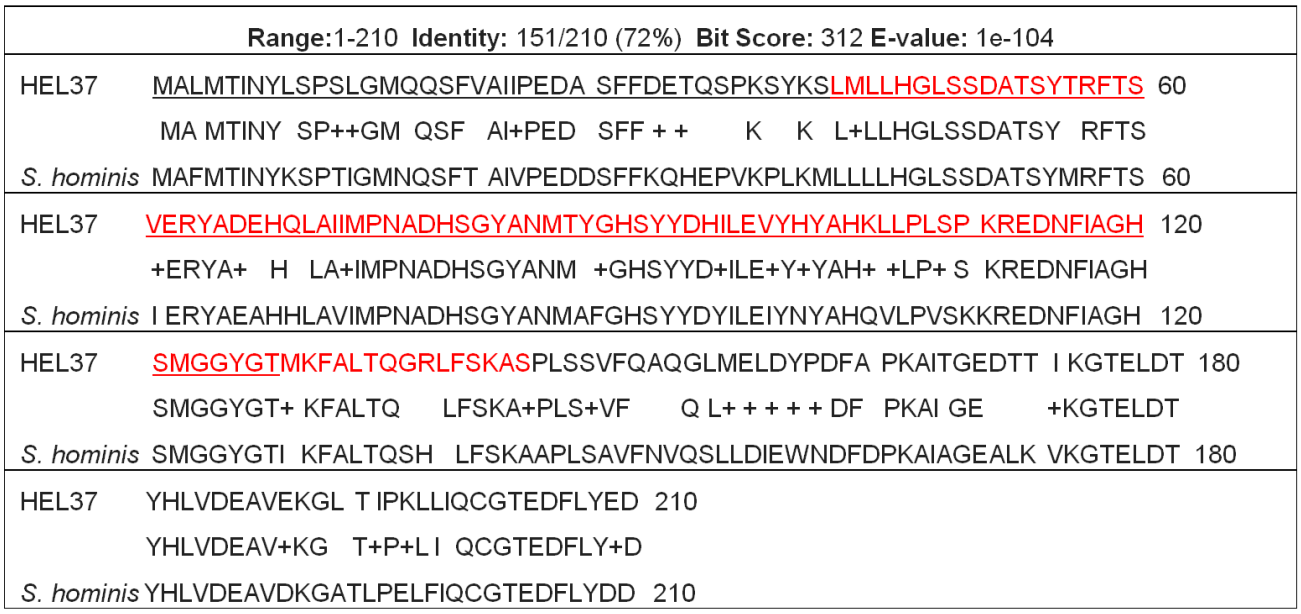

Figure 3.9. Alignment of HEL37 amino acid sequence to tributyrin esterase from *Staphylococcus hominis*. Protein sequence similarity of 72% was reported to this tributyrin esterase protein. Section in red of HEL37 indicates alpha/beta hydrolase family conserved domain COG2819 (Figure 3.7). Highlighted areas of selection indicate predicted esterase domains COG0627 (Figure 3.8).

NCBI BLAST of the HEL12 protein (247aa) found it to be most closely related to a hypothetical protein from *Citrobacter spp MGH55*, with an alpha/beta hydrolase family (pfam12695) conserved domain at the interval 29-229aa (Figure 3.9) (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml). Large inserts within this multi-domain family is found between amino acids 48-61, 112-143, and small amino acid insertions at 185-187 and 195-196.

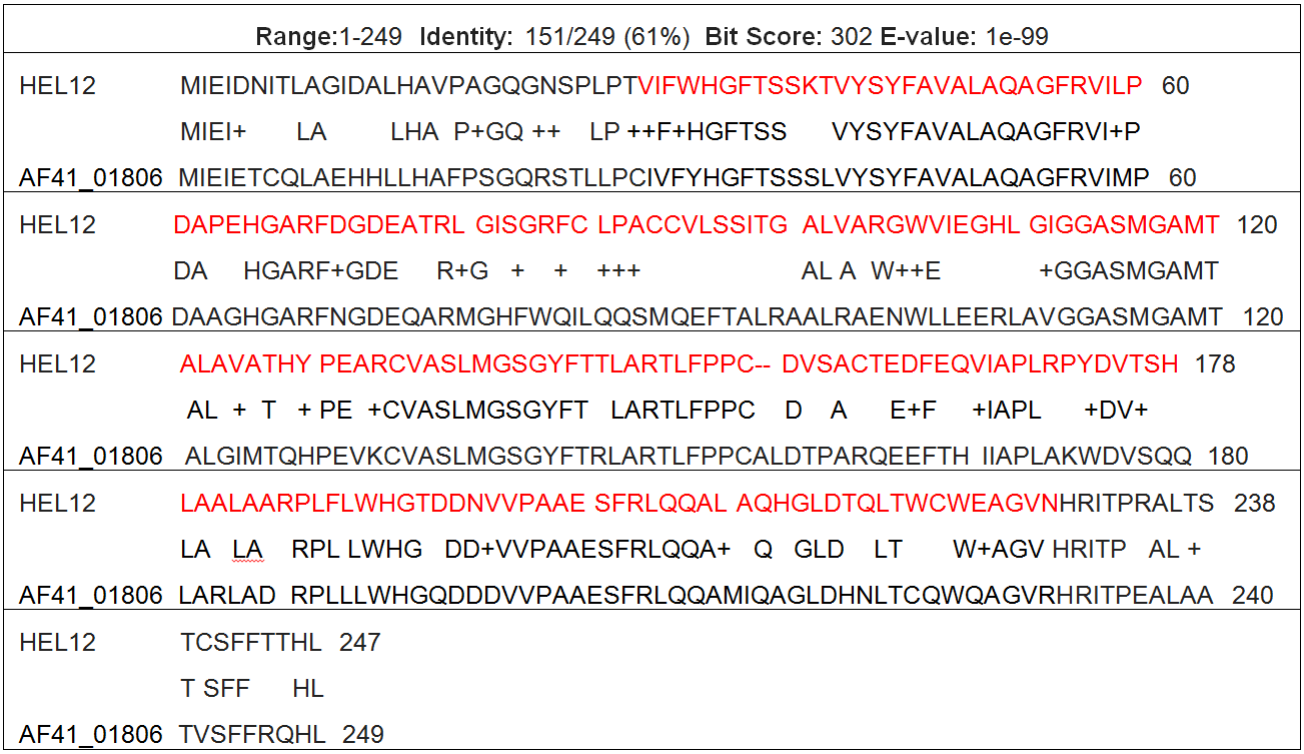

Figure 3.10. Alignment of HEL12 to Hypothetical protein AF41_01806 from *Citrobacter spp MGH55*. Sequence similarity of 61% was reported to this uncharacterised protein. Section in red of HEL12 indicates alpha/beta hydrolase family conserved domain (Figure 3.8). Various similarities of this domain can be seen to the hypothetical protein AF41_01806, concluding that it too indicates the presence of the same conserved domain.

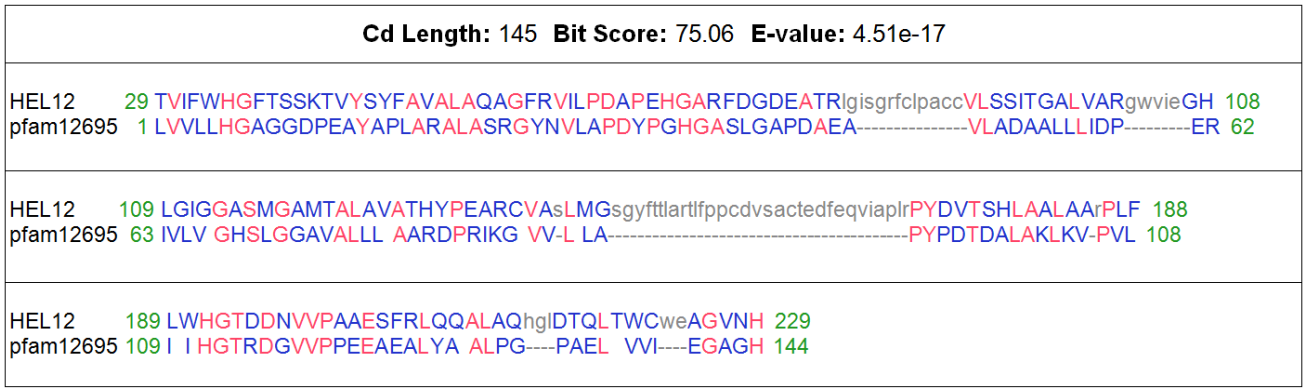

Figure 3.11. Alpha/beta hydrolase conserved domain pfam12695 aligned with amino acids 29-229 of HEL12. Amino acids highlighted in blue refer to amino acids that are similar but not identical. Those depicted in red refer to conserved amino acid residues.

3.2.2 Cloning of *HEL12* and *HEL37* into pRSET-C

Original bacterial genomic DNA extracts of the isolated bacteria were used to amplify *HEL12* and *HEL37* genes, respectively (Figure 3.12). The genes, *HEL12* and *HEL37*, were amplified via PCR from K4 and K6, respectively. These genes were then ligated into pGem-t-Easy for cloning and storage and then to pRSET-C and transformed into *E. coli* DH5-α.

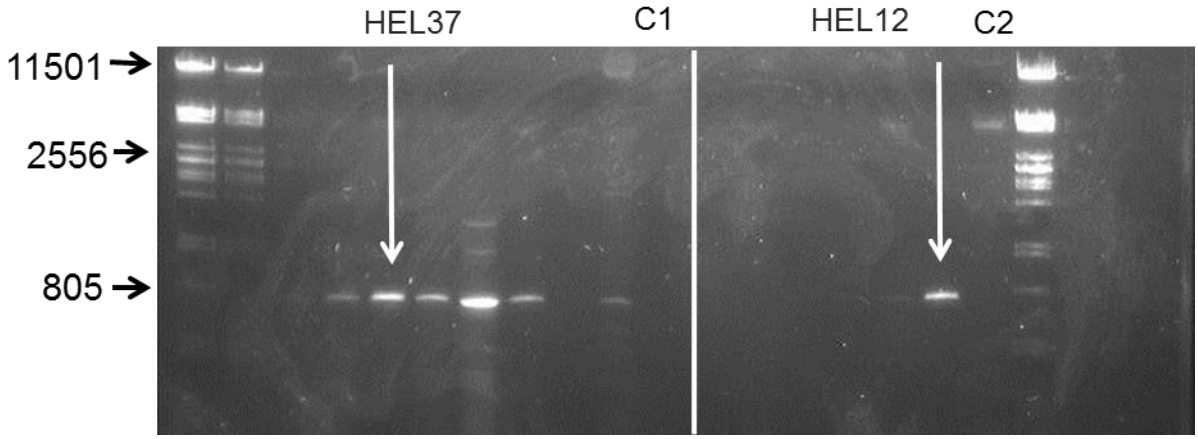

Figure 3.12. PCR amplification of *HEL37* and *HEL12* genes from native bacterial isolates. The PCR amplified fragments were cloned into pGem-t-Easy for further manipulation. In the left panel, the *HEL37* AE gene amplified from isolate K6 (identified as *B. vallismortis*) showed an approximate size of 800 bp. The right panel shows the amplification of *HEL12* from isolate K4 (identified as *C. farmeri*) at approximately 750 bp. C1 and C2 were included in the PCR reactions as respective negative controls.

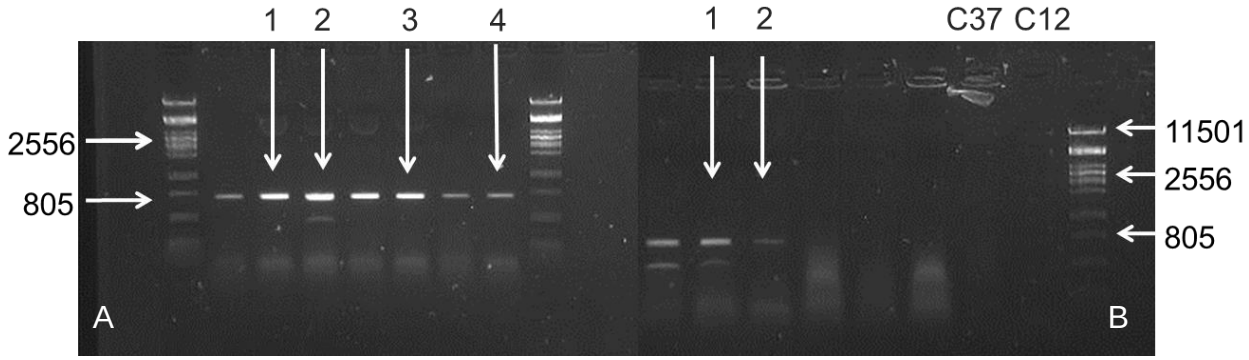

Figure 3.13. (A) Colony PCR confirmation of pGem-T-Easy:*HEL12* construct transformed into *E. coli* DH5-α. Amplification of 750 bp fragments was taken as positive clones. (B) Colony PCR confirmation of pGem-t-Easy:*HEL37* construct transformed into *E. coli* DH5-α. Amplification of 800 bp fragments was taken as positive clones.

HEL12 and *HEL37* containing plasmids were extracted and sequenced. Sequence analysis of these clones showed no mutations. The fragments, were excised with restriction enzymes *Bgl*II and *Hin*dIII, to retrieve gene and *HEL12* (approx 750bp), and *HEL37* (approx.

800 bp) (Figure 3.14). The genes were individually cloned into the expression vector pRSET-C.

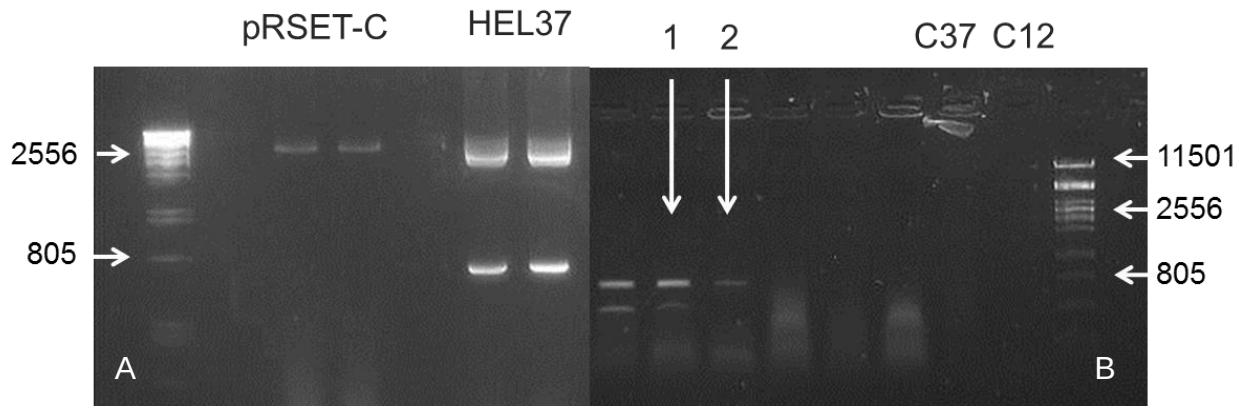

Figure 3.14. (A) Restriction digest (*Bgl*II and *Hind*III) of pGem-T-Easy:*HEL12*. (1) pGem-T-Easy:*HEL12* undigested (2) pGem-T-Easy:*HEL12* digested with *Bgl*II and *Hindl*II for excision of the 750bp fragment for cloning into pRSET-C. (B) Restriction digestion of pGem-T-Easy:*HEL37*. pGem-T-Easy:*HEL37* was restriction digested with *Bgl*II and *Hind*III, respectively to allow cloning into the expression pRSET-C.

3.2.3 Expression and purification of HEL12 and HEL37

The pRSET-C constructs containing the *HEL12* and *HEL37* encoding sequences were transformed into *E. coli* BL21, respectively. Pilot acetyl esterase activity tests were performed using crude extracts from *HEL12* and HEL37 overexpressing *E. coli* BL21 cell lines (Figure 3.15). The screen confirmed that when crude protein extracts from pRSET-C:*HEL12* and pRSET-C:*HEL37* overexpressing cell lines were presented with 2mM p-nitrophenyl acetate (in potassium phosphate, pH 7.4 buffer), a visible yellow colour indicated cleavage of this substrate. In comparison, the control (*E. coli* BL21 transformed with pRSET-C empty vector) did not show a colour change.

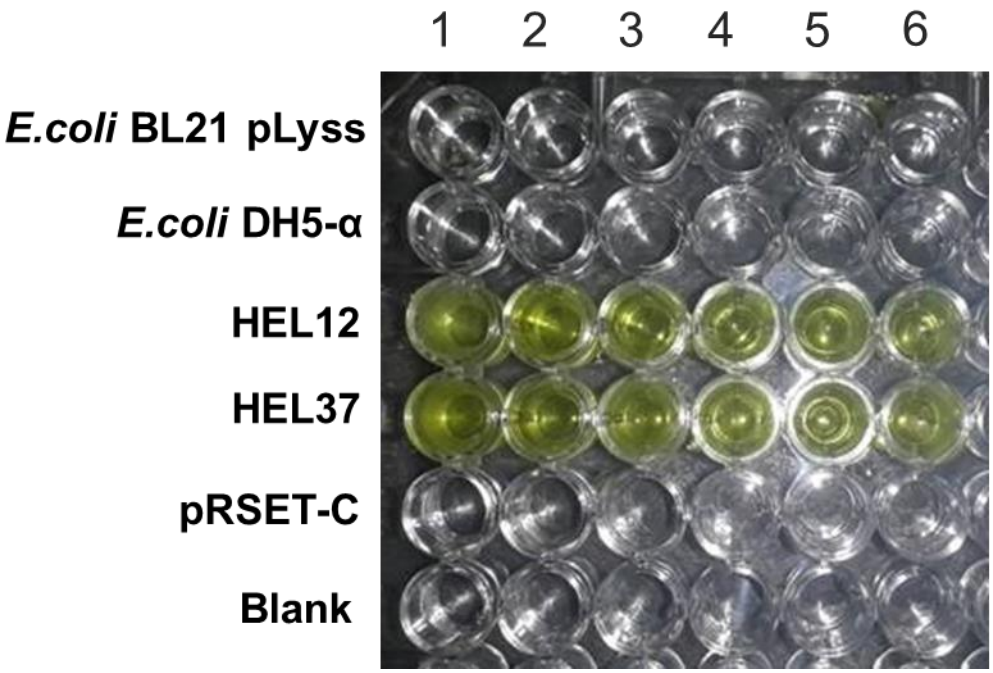

Figure 3.15. Pilot AE screen using crude protein extracts from HEL12 and HEL37 overexpressing BL21 cell lines. Columns 1-6 represent decreasing concentrations of p-nitrophenyl acetate (10 mM, 8 mM, 6 mM, 4 mM, 2 mM and 1 mM). *E. coli* BL21 pLyss, *E. coli* DH5-α crude and *E. coli* BL21 pLyss transformed with empty pRSET-C was used as controls. No colour reaction was observed for crude protein extracts from the controls. *E. coli* BL21 pLyss containing pRSET-C:*HEL12* and pRSET-C:*HEL37* showed a distinct yellow discolouration indicative of p-nitrophenyl acetate hydrolysis.

After pilot experiments showed that the *HEL12* and *HEL37* genes were expressed and showed activity on p-nitrophenyl acetate, it was decided to purify these enzymes. The individual enzymes were purified using Histidine-tag-Nickel affinity chromatography. Purified protein fractions were visualized by SDS-PAGE which estimated the size of the individual enzymes as approximately 30 kDa (Figure 3.16).

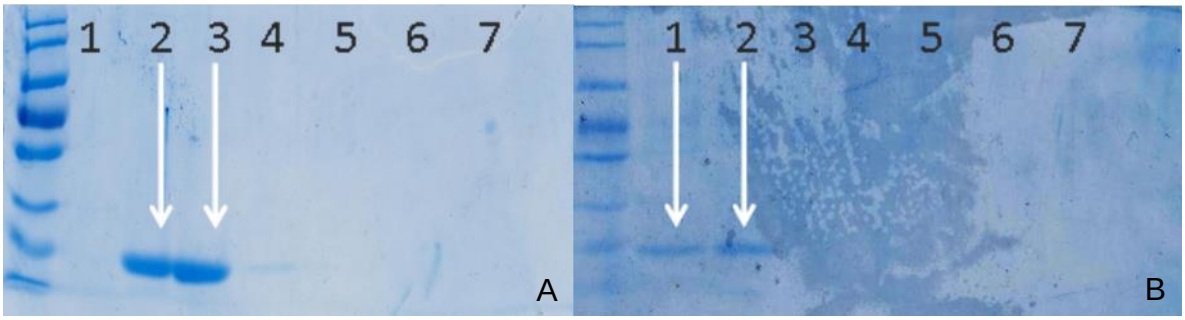

Figure 3.16. SDS-PAGE of (A) purified 6xHis-HEL12 and (B) purified 6xHis-HEL37 protein. Expression was induced with 1mM IPTG of pRSET-C:*HEL37* and pRSET-C:*HEL12* containing *E. coli* BL21 pLyss cultures. Crude protein preparations were subjected to Ni_2^+ affinity chromatography and fractions containing purified proteins were separated on 8% (v/v) SDS-PAGE.

3.2.4 Optimum Protein concentration of HEL12 and HEL37

Optimum protein concentration for assays was determined using standard AE assay conditions as determined in section 2.8.3. Optimum concentration was determined to be 10 ng/ml (Figure 3.17).

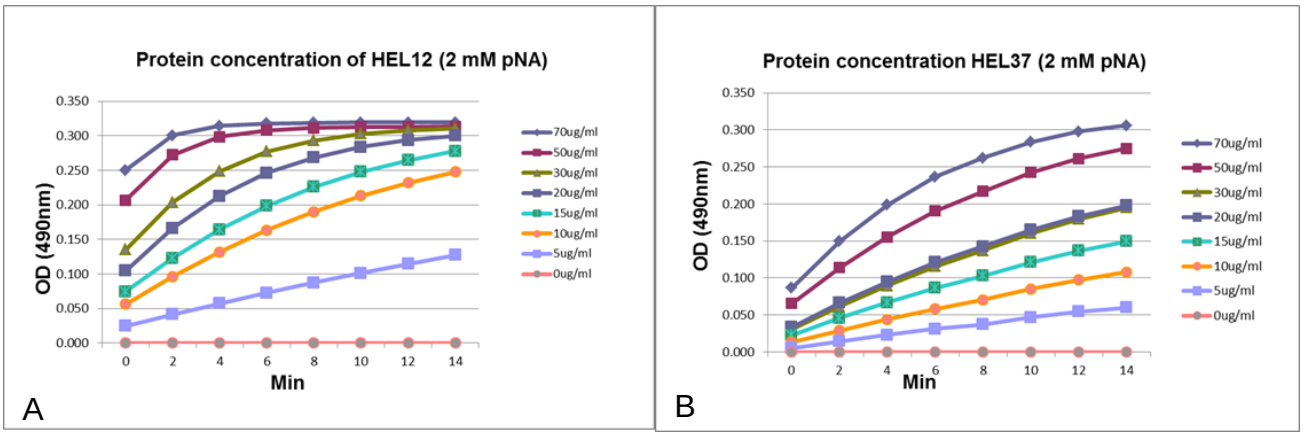

Figure 3.17. (A) Assessment of AE of the HEL12 protein with increasing protein concentration. Increasing AE activity assay of recombinant HEL12 protein was incubated with 2mM p-nitrophenyl acetate and monitored over 15 min. A logarithmic correlation was observed between increased concentration of enzyme and activity over time. Control sample had no protein extract. (B) Assessment of AE activity with increasing protein concentration. Assay of recombinant. Increasing amounts of purified HEL37 enzyme was incubated with 2mM p-nitrophenyl acetate and activity monitored over 15 min. A logarithmic correlation was observed between increased concentration of enzyme and activity over time. Control experiments contain no protein extract. The data points presented represent a mean ± S.E. of 8 repeats and two biological repeats.

3.2.5 Determining conditions for Standard Kinetic assay

Different buffer systems were evaluated for AE assays (Figure 3.18). Potassium phosphate (100 mM) at pH 9.0 (column 1) had a decreased degradation of p-nitrophenyl acetate when compared to sodium phosphate, pH 9.0 (column 2). Potassium phosphate buffer (Column 3) had less of an influence on p-nitrophenyl acetate at pH 7.2 than sodium phosphate (Column 4). Almost no degradation of the substrate was found when it was dissolved in potassium phosphate (pH 4.7); however sodium phosphate shows a significant amount of p-nitrophenol being present in the sample after 10 min of incubation.

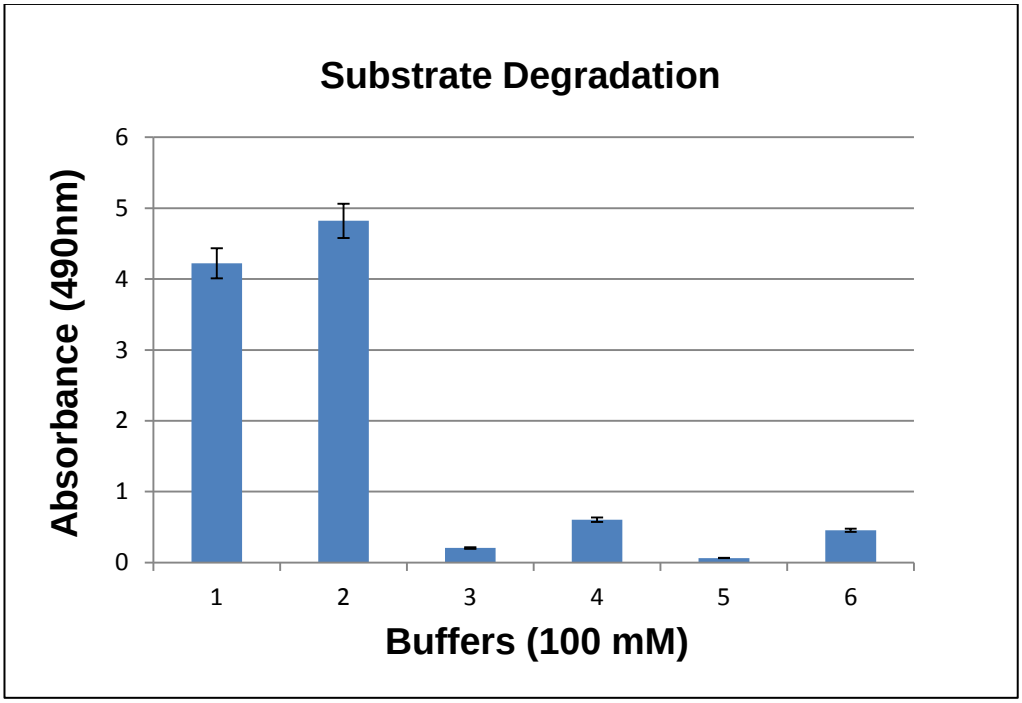

Figure 3.18. Effect of buffer ions and pH on the degradation of the substrate p-nitrophenyl acetate. Spontaneous substrate degradation was assessed in two different buffers and three buffer ranges. (1) Potassium Phosphate at pH 9.0, (2) sodium phosphate at pH 9.0, (3) Potassium Phosphate at pH 7.2, (4) sodium phosphate at pH 7.2, (5) Potassium Phosphate at pH 4.7, (6) Sodium phosphate at pH 4.7. Each data point presented in experimental results represents ±S.E.M of 8 independent replicates.

The AE substrate p-nitrophenyl acetate is known to have a short half-life when diluted with certain diluents (Levisson *et al.,* 2012). To test substrate degradation substrate stocks were monitored over a period of 4 days. Substrate (p-nitrophenyl acetate) dissolved in 100% ethanol at a concentration of 100 mM was diluted to 2 mM using 100 mM potassium phosphate and measurements were taken spectrophotometrically at 490 nm.

Spectrophotometric readings showed no significant deterioration when compared to same dilutions measured on the first day (Figure 3.19).

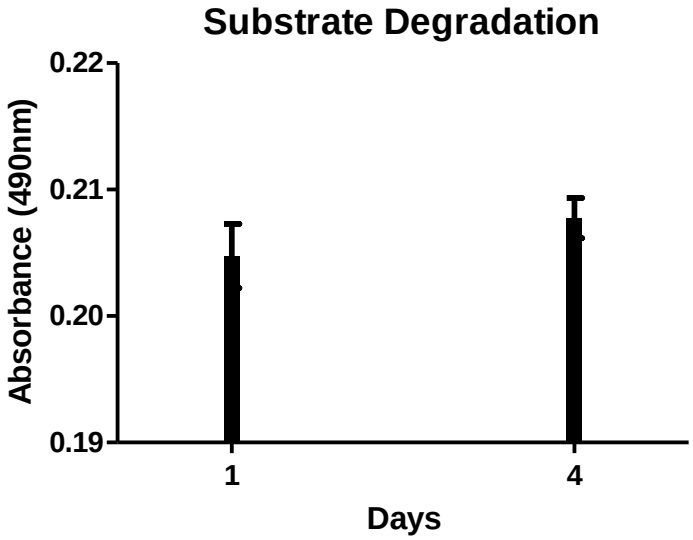

Figure 3.19. Degradation of p-nitrophenyl acetate over time. Substrate (p-nitrophenyl acetate) was diluted to 2mM using 100mM Potassium phosphate and measurements were taken spectrophotometrically at 490nm. Absorbance values were taken at day 1 and then on day 4. Results indicated no significant degradation of p-nitrophenyl acetate over a four day period.

3.2.6 Substrate specificity assay for HEL12 and HEL37

HEL12 and HEL37 recombinant proteins were assessed for their ability to hydrolyse various substrates. Assays showed that HEL12 had the most activity on p-nitrophenyl acetate, with an activity of 0.00372 mMol/min/mg purified protein, and HEL37 with an activity of 0.002 mMol/min/mg purified protein. HEL12 showed activity (0.00011 mMol/min/mg purified protein) on p-nitrophenyl butyrate, but no other selected substrate (Figure 3.20). HEL37 was unable to hydrolyse p-nitrophenyl butyrate, but small amount of activity (0.00017 mMol/min/mg protein) was detected for HEL37 on the substrate p-nitrophenyl palmitate (6), but no other selected substrate.

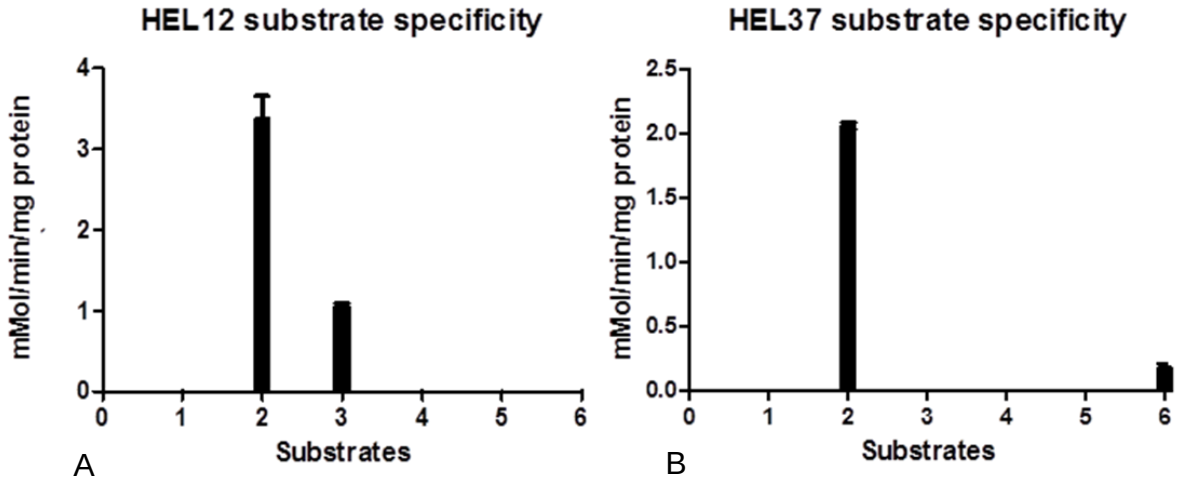

Figure 3.20. (A) Analysis of degradation of esterified substrates by the recombinant 6xHis-HEL12. The various p-nitrophenyl compounds used to test substrate specificity was diluted to 2 mM concentrations. (1) Control, (2) activity of 0.00372 Mol/min/mg of protein on p-nitrophenyl acetate, less activity (0.00011 Mol/min/mg protein) on p-nitrophenyl butyrate (3) and no activity on p-nitrophenyl octanoate (4), p-nitrophenyl dodecanoate (5) and p-nitrophenyl palmitate (6). (B), HEL37 showed activity on p-nitrophenyl acetate (0.00206 mMol/min/mg protein) and slight activity of 0.00017 Mol/min/mg protein on screening compound p-nitrophenyl palmitate.

3.2.7 pH Optima of HEL12 and HEL37

AE assays were conducted at various pH ranges from pH3.0-10.0. Activity assays indicated significant AE activity between the ranges of pH 7.0 to 8.0 for recombinant protein HEL12 and between pH 6.5 to 8.0 for recombinant protein HEL37 (Figure 3.21). The assays indicated significant AE activity for both enzymes at pH 8.0. In the case of both enzymes, no further activity was observed after surpassing the pH optimum was reached and it steeply declined.

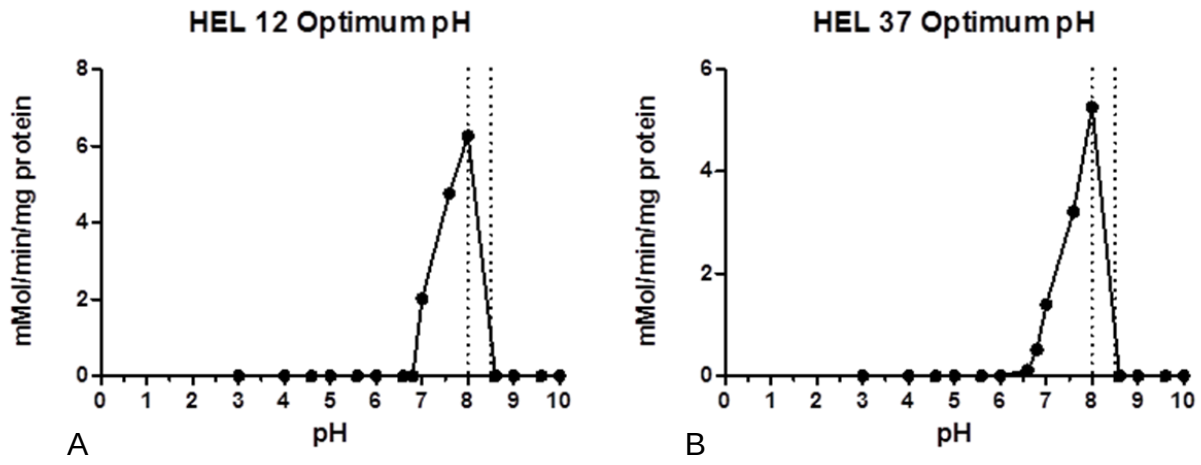

Figure 3.21. Effect of pH change on (A) 6xHis:HEL12 activity (B) 6xHis:HEL37 activity. Activity of enzyme was determined spectrophotometrically (OD_{490}) after 10 min at pH ranges from 3.0 to 10.0. Highest activity was detected at pH of 8.0. Each data point presented in experimental results represents ±S.E.M of 8 independent replicates and baseline normalization. The data points presented represent a mean ± S.E. of 8 repeats and two biological repeats.

3.2.8 Temperature Optima of HEL12 and HEL37

Enzyme activity was assessed at various temperatures (25 - 70°C). Within the range of temperatures that were assayed, the optimum temperature for recombinant protein HEL12 showed maximum activity at 30°C and a decreasing activity up to 50°C. A steep decline in activity was soon observed (55°C) and could not be detected after 60°C (Figure 3.22). In contrast, the recombinant HEL37 was found to be 35°C. However, we found a steep decrease in activity leading up to 60°C (Figure 3.22). Activity was drastically decreased at 65°C and no activity was found at 70°C and 80^{o}C.

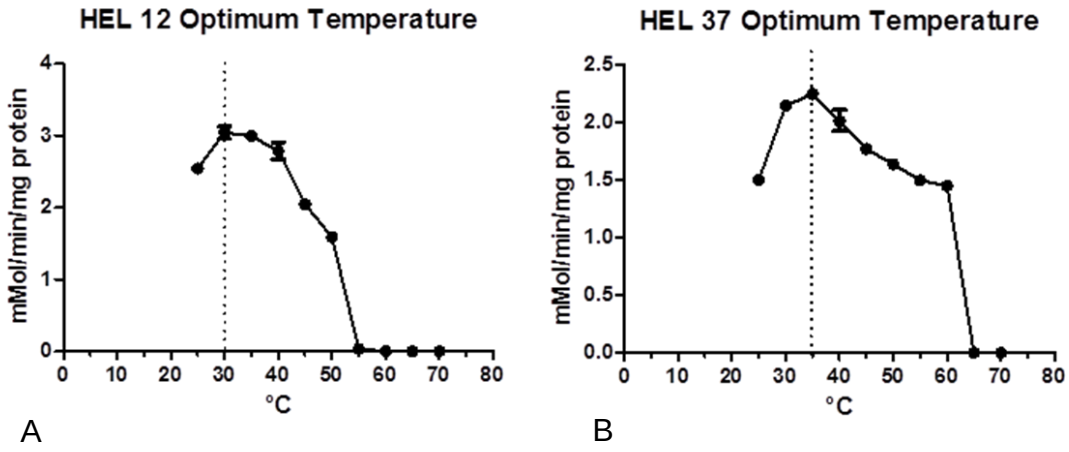

Figure 3.22. (A) Increasing temperature AE assay of recombinant protein HEL12. Assay was conducted from 25 to 70°C for 10 min. The highest activity peak was detected (OD_{490}) at 30°C indicating optimum enzyme activity. A sudden drop at 55°C indicated loss of activity and enzyme degradation. The data points presented represent a mean ± S.E. of 8 repeats and two biological repeats. (B) Temperature AE assay of recombinant protein HEL37. The highest activity was detected (OD_{490}) at 35°C indicating optimum enzyme activity. Each data point presented in experimental results represents ±S.E.M of 8 independent replicates and baseline normalization.

3.2.9 Kinetic properties (K_m and V_{max}) of HEL12 and HEL37

Kinetic parameters for the recombinant proteins HEL12 and HEL37 were determined for the substrates 4-nitophenyl acetate and p-nitrophenyl palmitate, respectively (**Figure** 3.23 and **Figure** 3.24). Parameters were determined using increasing amounts of substrate (0-14mM) under standard experimental conditions of 30°C and pH 7.0 (see section 2.8.2). The fastest reaction was observed by the recombinant protein HEL12 on p-nitrophenyl acetate with a K_m value of 1.321mM and a V_{max} = 3.812 mMol/min/mg protein (Figure 3.23). Recombinant protein HEL12 was also assessed on the substrate 4-nitophenyl butyrate with an observable V_{max} = 1.523 mMol/min/mg protein and a K_m = 1.692 mM (Figure 3.23). The 6xHis:HEL37 (Figure 3.24) had an observed V_{max} = 2.462 mMol/min/mg protein and a K_m = 1.621 mM on the substrate p-nitrophenyl acetate. Protein 6xHisHEL37 (Figure 3.24) had an observed V_{max} = 0.4363 mMol/min/mg protein and a K_m value of 3.571 on the substrate p-nitrophenyl palmitate. Substrate concentrations with respect to the kinetic assays where shown to be optimal at a concentration of 10 mM.

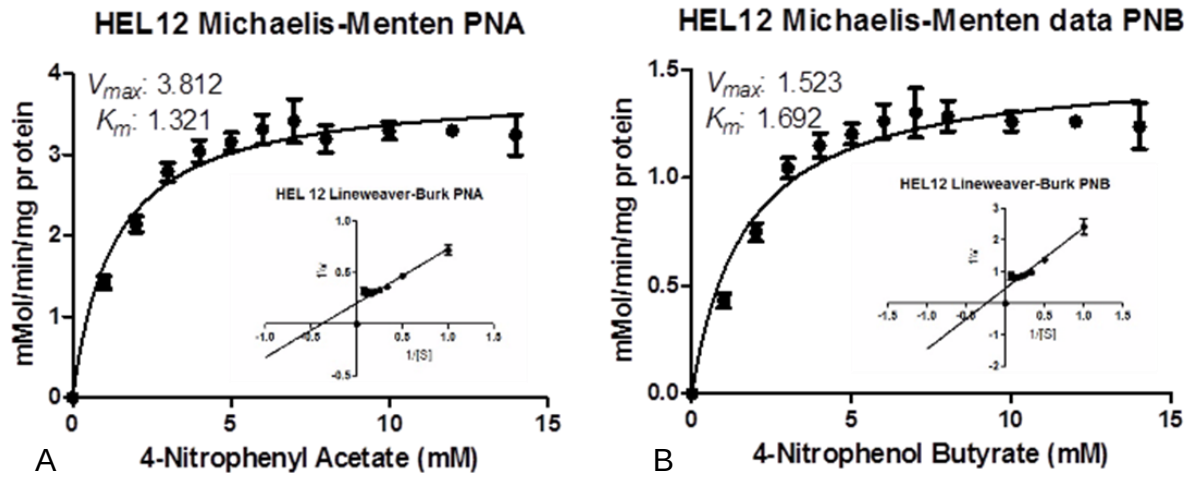

Figure 3.23. (A) Michaelis-Menten with inset Lineweaver-Burk plot of kinetic parameters of 6xHis:HEL12. Assays were performed using the substrate p-nitrophenyl acetate substrate at a various concentrations ranging from 0 to 14mM. The calculated K_m = 3.812mM and V_{max} = 1.321 mMol/min/mg protein. (B) Kinetic parameters determined for 6xHis:HEL12. Data represents Michaelis-Menten with inset Lineweaver-Burk plot using the substrate p-nitrophenyl butyrate. The determined parameters (K_m = 1.692 mM and V_{max} = 1.523 mMol/min/mg protein). Standard AE reaction conditions were used as described earlier (section 2.8.2). Each experimental data point represents ±S.E.M of 8 independent replicates and baseline normalization.

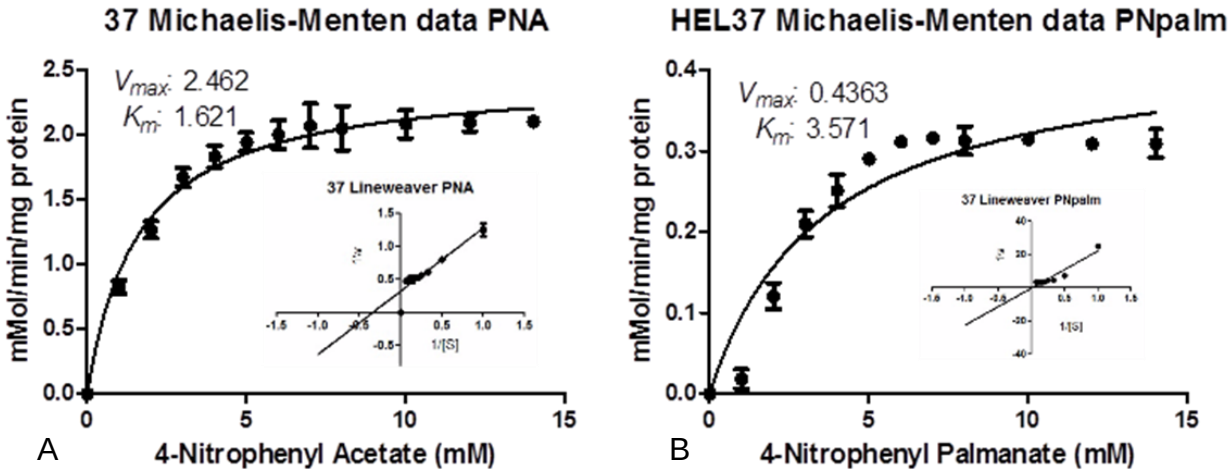

Figure 3.24. (A) Michaelis-Menten plot with inset Lineweaver-Burk plot for 6xHis:HEL37. Assays were performed using p-nitrophenyl acetate substrate at a concentration range of 0-14 mM. The calculated was K_m = 1.621 mM and V_{max} = 2.462 mMol/min/mg protein was determined using standard Michaelis-Menten calculations. (B) Michaelis-Menten plot with inset Lineweaver-Burk plot for 6xHis:HEL37. Assays were performed using p-nitrophenyl palmitate substrate at a concentration range of 0 to 14 mM. The kinetic parameters were calculated to be K_m = 3.571 mM and V_{max} = 0.4363 mMol/min/mg protein. Standard AE reaction conditions were used as described earlier (section 2.8.2). Each point in the data set represents ±S.E.M of 8 independent replicates and baseline normalization.

3.3.0 Native Substrate Assays

Native substrate specificity and activity for recombinant proteins HEL12 and HEL37 were determined by incubating respective enzymes with seven substrates using protocols described in section 2.8.8. Recombinant protein 6xHis HEL12 (Figure 3.25) had detectable acetic acid release on various substrates: (1) wheat arabinoxylan (0.193 mMol/min/mg protein), (2) xylan from Oats (0.573 mMol/min/mg protein), (3) arabinan from sugar beet (0.010 mMol/min/mg protein), (5) sucrose acetate isobutyrate (0.314 mMol/min/mg protein), (6) chitin from shrimp (0.190 mMol/min/mg protein) and (7) xylan from beechwood (0.620 mMol/min/mg protein). The 6xHis HEL37 (Figure 3.25) showed detectable acetic acid release on a few possible native substrates: (1) Wheat arabinoxylan (0.059 mMol/min/mg protein), (2) Xylan from Oats (0.292 mMol/min/mg protein), (6) Chitin from shrimp (0.095 mMol/min/mg protein) and (7) Xylan from Beechwood (0.310 mMol/min/mg protein).

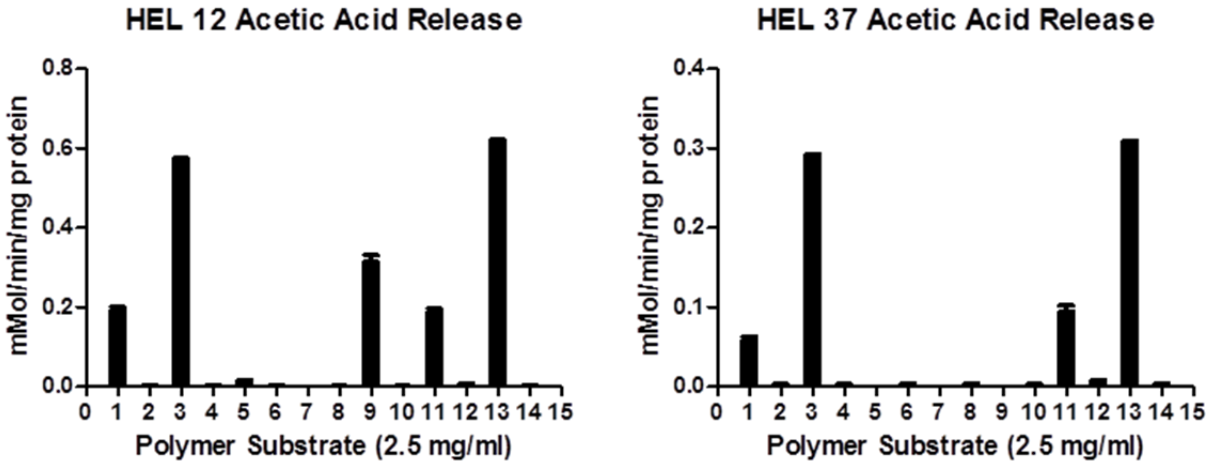

Figure 3.25. Analysis of native polymer substrate hydrolysing capability of (A) recombinant 6xHis-HEL12 and (B) recombinant 6xHis-HEL37 on plant cell wall derived substrates. Various esterified compounds were used to test substrate specificity of HEL12 and HEL37 enzymes. All substrates were diluted to 2.5 mg/ml concentrations in potassium phosphate. Each substrate was assessed both with purified protein extract and *E.coli* extract not containing the vector (1,2) wheat arabinoxylan (Megazyme), (3.4) xylan from Oats (Sigma-Aldrich), (5,6) arabinan from sugar beet (Megazyme), (7,8) xyloglucan from tamarind seed (Megazyme), (9,10) sucrose acetate isobutyrate (Sigma-Aldrich), (11,12) chitin from shrimp (Sigma-Aldrich), (13,14) xylan from beechwood (Sigma-Aldrich), (15) baseline control.

The 40kDa *Bacillus pumilus* PS213 enzyme was screened for activity by Degrassi *et al.,* 1998 only on one substrate, p-nitrophenyl acetate and recorded to have K_m and V_{max} values of 1.54 mM and 0.360 mMol/min/mg respectively, with an optimum temperature of 55°C and pH of 8.0. The 40kDa *B. pumilus* AE enzyme was recorded to have a K_m and T_m values of 0.50 mM and 0.126 mMol/min/mg respectively, an optimum pH of 8.5 and a temperature of 45°C (Table 3.5).

Table 3.5. HEL12 and HEL37 protein parameters are related to known esterase enzymes. All of the related acetyl xylan esterases were found to have neutral to basic pH ranges. Sizes of the esterases were found to vary considerably, as was the optimum temperatures. The enzymes where primarily all assessed on p-nitrophenyl acetate.

Protein	Primary Substrate	K_m (mM)	V_{max} (mmol/min/mg)	Reference
HEL12	**p-nitrophenyl acetate**	**1.612**	**2.462**	
HEL37	**p-nitrophenyl acetate**	**3.812**	**1.321**	
B. pumilus PS213	p-nitrophenyl acetate	1.54	0.360	Degrassi *et al.,* 1998
TM0077	p-nitrophenyl acetate	0.185	0.145	Levisson *et al.,* 2012
***B. pumilus* AXE**	p-nitrophenyl acetate	0.50	0.126	Dengrassi *et al*., 2000

Protein	Secondary substrate	K_m (mM)	V_{max} (mmol/min/mg)	Reference
HEL12	**p-nitrophenyl palmatate**	**3.571**	**0.436**	
HEL37	**p-nitrophenyl butyrate**	**1.692**	**1.523**	
B. pumilus PS213				Degrassi *et al.,* 1998
***B. pumilus* AXE**	4-methylumbelliferyl acetate	0.50	0.98	Dengrassi *et al*., 2000
TM0077	p-nitrophenyl propanoate	0.137	0.125	Levisson *et al.,* 2012

Protein	Optimum pH	Optimum Temp	Size	Reference
HEL12	**8.0**	**35°C**	**33 kDa**	
HEL37	**8.0**	**30°C**	**25 kDa**	
B. pumilus PS213	8.0	55°C	40 kDa	Degrassi *et al.,* 1998
***B. pumilus* AXE**	8.5	45°C	40 kDa	Dengrassi *et al*., 2000
TM0077	7.5	100°C	37 kDa	Levisson *et al.,* 2012

3.2.10 Cloning of *HEL12* and *HEL37* into pCambia 1300:BinAR

In order to target expression of *HEL12* and *HEL37* proteins to the Golgi apparatus the 36 amino acid signal sequence, *A. thaliana* β1,2-Xylosyltransferase (*Xyl-T36*), was amplified from *Arabidopsis thaliana* and cloned into pCambia1300 plant vector (Figure 3.36).

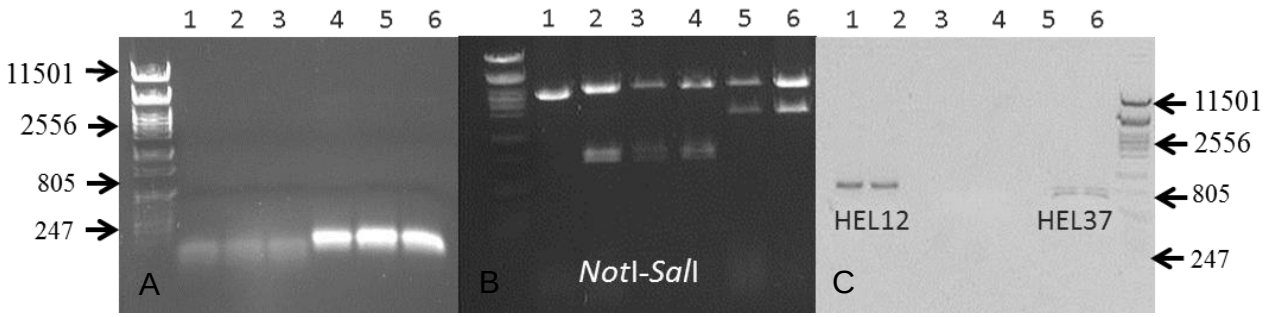

Figure 3.36. (A) PCR amplification of *Xyl-T36* from Arabidopsis cDNA, (1) negative water control, (2) empty plasmid control with primers, (C 3) *E. coli* crude genomic DNA extract. (1, 2 and 3) *Xyl-T36* amplicon situated at 108 bp in a 2% agarose gel. (B) *Not*I-*Sal*I Restriction digest of fragments in pGem-T-Easy confirming recombinant protein for plant expression. (1) *Xyl-T36::HEL12* clone 1 with no insert, (2) *Xyl-T36::HEL12* clone 2 showing excision of insert (Approx. 900 bp), (3) *Xyl-T36::HEL37* Clone 1 showing excision of insert (Approx. 900bp), (4) *Xyl-T36::HEL37* clone 2 showing excision of insert (Approx. 900bp). (5 and 6) is uncut empty vector. (C) PCR amplification of pCambia 1300:*Xyl-T36::HEL12* and pCambia 1300:*Xyl-T36::HEL37*. A negative water control (C1) and negative pCambia 1300 empty vector was used to substantiate the amplification of the fragments.

Chapter 4

4. Discussion

In higher plant-derived biomass esterified bonds are often found on the hemicellulose (xylan) backbone and are subject to hydrolysis by AE enzymes. Recently hydrolysis of hemicellulose polysaccharides have gained attention as an alternative renewable energy source due to rising fossil fuel costs. In this study, ten cell wall degrading bacteria were isolated and identified from a sample of saw dust (Cape Saw Mill, Stellenbosch, South Africa). Samples were collected on the rationale that an environment high in plant cell wall polysaccharides, microorganisms will tend to utilise these polysaccharides as energy sources. Bacterial isolates were identified based on lypolitic analysis and assessed for their esterase/lipase and acetyl esterase activity (Table 3.1). All of the strains produced plant cell wall degrading enzymes; however activity was only observed in the presence of lignocellulosic or synthetic compounds. Presence of enzyme activity only with the presence of specific substrates suggests that their production is inducible by and dependant on these specified substrates. A small insert library was constructed using genomic DNA isolated from these ten isolates. Through functional screening of the library clones for general esterase/lipase activity and further acetyl esterase activity, four candidate clones with high esterase activity were identified. Further characterization resulted in narrowing the pool to two clones, HEL12 and HEL37, which was biochemically characterized, and assessed for acetyl xylan esterase activity.

4.1 Bacterial classification

After initial selection of isolates, bacteria were classified using standard microbial techniques and 16S RNA genotyping. Diverse sampling conditions showed a mixture of gram positive and gram negative staining bacteria. All strains were found to be mesophillic, however their growth conditions ranged from aerobic to facultative anaerobic. Limitations of incubating the crude dilutions of the samples at 37°C in an aerated incubator contributed to the isolated bacteria being mesophillic and aerotolerant as opposed to the desired thermophilic isolates. Future studies should investigate high heat incubation of collected samples and should also include incubation conditions for the isolation of anaerobic microorganisms.

Phylogenetic analysis, of the 16S RNA sequences of the ten isolated bacteria has shown that they clustered well with the classic Neighbour-Joining/Maximum Likelihood tree construction (Figure 3.2). Analysis correlates the isolates to species and genera of ATCC curated taxa (Figure 3.2). It revealed a diverse ecology of AE producing bacteria from the Cape Saw Mill in Stellenbosch. As the environment was rich in plant biomass, a diverse nature of the isolated bacteria was expected. This indicates that the ability to break down plant cell wall polysaccharides as a carbon source extends to many different species of bacteria.

The classification studies enabled the identification of novel AE activites. Literature review highlighted no lipase genes discovered in nearest relatives to C1 (*Staphyococcus sp. IARI-S*-21), C7 (*Acinetobacter sp. LC084*) and K4 (*Citrobacter farmerii*) (Table 3.3). Similarly, no literature on esterase activity was found for the nearest relatives of C3 (*Staphylococcus epidermis*) and K4 (*Citrobacter farmerii*) (Table 3.2; 3.3 and 3.4). As stated previously, in our survey several isolates showed esterase/lipase activity (Table 3.1), although some of these isolates (C1, C7, and K4), in addition to esterase/lipase activity also showed acetyl esterase activity (Table 3.1). This alluded to possible novel acetyl xylan esterase enzymes present in C3 and K4 and novel lipase enzymes present in C1, C7 and K4. In an environment like this, a competitive factor exists as to which organisms can digest the carbon sources most effectively, and also which can utilise the more complex polysaccharides. It is due to this, that bacteria C3 and K4 adapted these specialised highly effective enzymes to utilise the xylan polymers present in the cell walls faster and more efficiently than their competitors.

Genus *Staphylococcus* [isolates C1 (*Staphylococcus sp. IARI-S*-21) and C3 (*Staphylococcus epidermis*)] have no current described enzyme considered to have lipase activity, but esterase and cellulase enzyme activity are common for the genus (Hasan *et al.,* 2006; Takahashi *et al.,* 1999). *Bacilli* [isolates *B. pumilus* (C7)*, B. thuringiensis* (C10)*, B. jetogali* (K8) and *B. valismortis* (K6)] are one of the most common bacteria found in environmental samples (Alcaraz *et al.,* 2010; Yakimov *et al.,* 1995). The extracellular compounds and enzymes from *Bacilli* species have been used in various industrial processes, including polysaccharide degrading enzymes, cellulases, lipases and esterases for the production of biofuels (Zhang *et al.,* 2010a). *Acinetobacter* [isolate C8 *(A. lwoffii*)] is a species of bacteria commonly found in soil and plant samples. *Acinetobacter* species have

well documented cellulase, lipase and esterase enzymes (Nair, 2007; Du *et al.,* 2010; Shebab *et al.,* 2010). The species is known for various industrially important enzymes. *Klebsiella* [isolate K2 *(K. pneumonia*)] species are gram negative, and the most common organism soil and water samples. Cellulase, lipase and esterase activity has been documented in various *Klebsiella* species, but are not currently exploited in industrial processes (Han, 1995; Wang and Tsai, 2009; Mu *et al.,* 2008). *Cronobacter* [isolate K3 (*C. sakazadkii*)] is a bacterium commonly found in nitrate rich environments for example soil and wood samples. *Cronobacter* species are known for acetoin production as well as other volatile hydrocarbons by breaking down cellulose, lipids and various esterified polysaccharides (Ma, 2006; Shabeb *et al.,* 2010).

Bacterial isolates C10, K2 and K3 revealed AE activity without showing any esterase/lipase activity. Assessment of AE enzymes from native isolates can be influenced by bacterial fermentation which would lower the pH of the solution and cause a greater degradation of screening substrate p-nitrophenyl acetate. This was overcome by buffering the crude protein extract to pH 7.4 with potassium phosphate buffer before the assays were performed. Following this protocol it would conclude that the results of this assay for C10, K2 and K3 were not influenced by fermentation of other simple sugars in the incubation broth. This indicates that the AE enzymes present in these isolates are not lypolitic, but specifically adapted to esterase activity on short carbon-chain molecules. It is possible that these esterases, although they can hydrolyse the synthetic compound p-nitrophenyl acetate are incapable of hydrolysis of longer chain polysaccharides.

Isolate K4 (*C. farmeri*) which yielded the isolation of the *HEL12* open reading frame, is a bacterium associated with a variety of medical ailments in hospitals where it is an opportunistic pathogen. Generally it is found in environments with high human occupation, it is associated with minor ailments and focusses on breakdown of human refuse, which often contains various cell wall polysaccharides. It is therefore, not surprising that it was found in a sample of saw dust present amongst saw mill employees, nor that it is associated with AE enzyme activity, although none has been classified to date. Being an *Enterobacter* of the phylum *Proteobacteria*, it is widely separated from K6 (*B. vallismortis*) which yielded *HEL37* of the phylum *Firmicutes*. *Bacilli* however, are found in all environments on earth. *B. vallismortis* is closely related to *B. subtilis* both of which are found in soil samples that are high in plant polysaccharides (Roberts *et al.,* 1996). These two isolates (viz. K4 and K6), even though they are widely removed from each other in the bacterial taxa, are not

uncommon to be found within the same vicinity where polysaccharide chains can be hydrolysed as primary energy source. It therefore stands to reason that both bacteria were found in the degraded saw dust sample to utilise the plant cell wall polysaccharides as an energy source. Whereby they would have to be hydrolysing various side-chains including acetyl groups to allow economic and complete usage of the polymeric backbones of the plant cell wall.

4.2 Identification of AE clones

Most studies that are based on library screening have a high number of positive clones, causing high-throughput screening to be required. Due to the isolation of 53 positive clones from the screened library, efficacy of each of these expressed active p-nitrophenyl acetate hydrolysing proteins needed to be assessed.

In this study we were screening for active AE enzymes. The initial screen for lypolitic activity identified 53 clones that formed the base of the study for more extensive scrutiny. This methodology enabled a screen focussing only on the clones that were of interest, streamlining the labour-intensive process as well as making it more cost-effective. By screening for lypolitic activity first, and disregarding any clones that does not have lypolytic activity, enzymes that are only p-nitrophenyl specific was excluded from this study, allowing focus on enzymes that could hydrolyse esterified polymers as well as lipids.

A preliminary AE activity screen was initially commenced by grouping each of the samples into 12 groups of 5 clones to save on substrate and allow for fast, effective screening of all 53 positive clones. Only groups with a high hydrolysis rate (2 mMol/min/mg) were selected for final individual analysis. A considerable negative aspect of this method of screening is that if only one of the five pooled clones were to have AE activity, and in another group four of the five would have activity it is plausible that one could potentially omit a highly active AE enzyme from further studies. However likely as this was, the aim of the study was to identify and characterize novel acetyl xylan esterase proteins, which would be achieved regardless. This allowed us to continue studies with five of the 12 groups for individual screening, further streamlining the process. By using high criteria (2 mMol/min/mg) in the final AE analysis the process allowed identification of only the most active clones.

4.3 Characterisation of HEL12 and HEL37 enzymes

DNA sequencing in conjunction with amino acid analysis by NCBI-BLAST algorithm is a powerful tool to allow the discovery of protein conserved domains and nearest relatives. Amino acid analysis highlighted HEL12 containing an alpha/beta hydrolase conserved domain (pfam12695). This domain is found in various esterases, thioesterase and conserved hypothetical proteins in *Clostridia* species. Domain pfam12695 is indicated in a group of hydrolases that are currently not classified in a specific super-family. Presence of this domain alluded that protein HEL12 is associated with ester bonding and hydrolysis. Domain pfam12695 also being associated with acyl transferase genes, which do not hydrolyse ester bonds, but only transfer them to a binding site could then imply that Domain pfam12695 has been misinterpreted as a hydrolytic domain, and rather be associated with the binding of acyl groups to hydrolytic or transportation enzymes (Riley *et al.,* 2006). The close relation of HEL12 with hypothetical protein AF41_01806 from *Citrobacter spp MGH55* also containing domain pfam12695 strongly indicates hypothetical protein AF41_01806 could be associated with acetyl group binding, or hydrolase activity. *Citrobacter* species are well known for their ability to break down cell walls, xylanases and cellulases have been identified in *Citrobacter* species present in the gut of termites (*Coptotermes formosanus)*. This would prompt that the hypothetical protein AF41_01806 could be a possible plant cell wall degrading enzyme (Mathew *et al.,* 2011; Okolie and Ugochukwu 1988).

Using the NCBI-BLAST algorithm HEL37 protein (210 aa) was found to be most closely related to tributyrin esterase protein from *S. hominis*. Alignment of this protein with tributyrin esterase sequence from *S. hominis* shows 72% similarity to *S. hominis* tributyrin esterase (Table 3.4). The indications of conserved multi-domain areas COG0627 and COG2819 being associated with hydrolase and esterase activity in various organisms shows that protein HEL37 can be associated with ester hydrolysis. The conserved area COG0627 (Figure 3.7) was found in S-formylglutathione hydrolase from *Agrobacterium larrymoorei* and *Salmonella entrica*. Conserved area COG0627 is not assigned to any specific domain superfamily. Alpha/beta hydrolase domain COG2819 (Figure 3.8) was found in hypothetical proteins in *Clostridiales bacterium* as well as putative esterase family protein species in *Clostridium sp. KLE1755* and *Roseburia intestinalis* (Weinstock *et al.,* 2013). Further analyses indicates deletions within the multi-domain alpha/beta hydrolase conserved region at interval 41-58aa. These modifications could give rise to differences in the binding domain of this enzyme ultimately affecting its affinity for the substrate. The association of these domains with hydrolytic activity and esterase family proteins indicate that HEL37 should

have esterase activity. Enzyme HEL37 could therefore be a dual functional enzyme, capable of tributyrin esterase and AE activity. Deletion points seen in HEL37 with regard to COG0627 and COG2819 would also indicate that the two conserved domains could be be part of a larger multi-domain family that allows for binding to a xylan polymer and facilitate ester hydrolysis from the backbone.

4.4 Kinetic association with other known esterases

The first isolated gene encoding for an AE was isolated from *Caldocellum saccharolyticum* (Lüthi *et al.,* 1990). Since then AE genes were isolated from various fungal and bacterial strains, most notably from *Aspergillus niger*, *Aspergillus aculeatus, Thermotoga maritima* and *Bacillus pumilus* (Levisson *et al.,* 2012; Degrassi *et al.,* 1998; Christgau *et al.,* 1995; de Graaff *et al.,* 1992). Characterization of their gene sequences has revealed separate xylan binding and catalytic domains (Ferreira *et al.,* 1993). In this regard protein HEL12 would have a separate hydrolytic domain associated with conserved domain pfam12695. HEL37 having two separate domains (viz. COG0627 and COG2819) does conform to literature regarding a separate xylan binding and hydrolytic domains.

HEL12 (33 kDa) and HEL37 (25 kDa) both function as AE enzymes, come from the same sample and are from bacteria distant from one another, yet are commonly found in the same niche. This leads to the conclusion that these enzymes would function in a similar manner. Both enzymes were smaller than literature comparisons; *B. pumilus* PS213, *B. pumilus* AE and *T. maritima* (Levisson *et al.,* 2012; Degrassi *et al.,* 2000; Degrassi *et al.,* 1998). It was found in this study that HEL12 had high activity on primary assay substrate p-nitrophenyl acetate (K_m = 3.81 mM and V_{max} = 1.321 mMol/min/mg protein), it had a low affinity for p-nitrophenyl acetate and also a slower maximum velocity for catalysis of the substrate than HEL37 (K_m = 1.612 mM and V_{max} = 2.462 mMol/min/mg protein). These presumably similar enzymes had a different substrate affinity, and also did not hydrolyse the same secondary substrates. HEL37 displayed a low affinity and low hydrolysis rate for p-nitrophenyl palmitate (K_m = 3.571 mM and V_{max} = 0.436 mMol/min/mg protein). HEL12 indicated a higher affinity for p-nitrophenyl butyrate than the primary substrate, and also had a higher velocity for the secondary substrate p-nitrophenyl butyrate. This indicates that HEL12 can be useful in an environment where short to intermediate carbon-chain side chain molecules exist, whereas HEL37 seems to primarily hydrolyse short chain O-acetyl groups, with a low affinity for a longer chain (C16) carbon chain molecule.

It is interesting to note that the slowest hydrolysis velocities of the isolated enzymes HEL12 and HEL37 would still be higher, or coincide with the primary substrate hydrolysis velocity of other documented enzymes assayed on the same substrate viz. *B. pumillus PS213* (K_m= 1.54 mM and V_{max} = 0.360 mMol/min/mg protein), *B. pumillus* AE (K_m= 0.50mM and V_{max} = 0.98 mMol/min/mg protein) and *T. maritima* AE enzyme TM0077 (K_m= 0.50mM and V_{max} = 0.126 mMol/min/mg protein) (Levisson *et al*., 2012; Degrassi *et al.,* 2000; Degrassi *et al.,* 1998). A phenomenon that could be explained by the stringent selection process during the screening for AE enzymes, that has resulted in isolation of enzymes able to hydrolyse p-nitrophenyl acetate with a high enough velocity to be recognised by the screening parameters.

While a desired thermophilic acetyl xylan esterase could not be identified in this study, the incubation of the samples at 37°C prior to microorganisms isolation selected for a general microbial diversity is conducive to mesophyllic bacteria. At an optimum temperature of 35°C and pH of 8.0 HEL12 is better suited to higher temperature hydrolysis than HEL37 displaying an optimum temperature of 30°C at a pH of 8.0. When this is compared to *B. pumillus PS213* (55°C at pH 8.0), *B. pumillus* AE (45°C at pH 8.5) and *T. maritima* AE enzyme TM0077 (100°C at pH 7.5) (Levisson *et al*., 2012; Degrassi *et al.,* 2000; Degrassi *et al.,* 1998) both enzymes fall within the parameters expected by documented literature of similar studies, but neither enzyme reach the same optimal temperature as any of the cited similar enzymes from literature. The reason for this could possibly be due to low incubation temperature of the original collected environmental sample.

We found HEL12 and HEL37 enzymes are smaller and more active compared to the AE enzymes isolated from *B. pumilus* and *T. maritima* (Levisson *et al*., 2012; Degrassi *et al.,* 2000; Degrassi *et al.,* 1998). Although their overall affinity for the substrate is lower, the velocity at which they hydrolyse p-nitrophenyl acetate makes them more industrious. Both HEL12 and HEL37 were shown to be significantly faster at hydrolysis than other classified AE enzymes, this fast hydrolysis rate is novel within this group of enzymes and is of industrial importance.

Since there are an abundance of screening substrates for characterization of these enzymes, p-nitrophenyl acetate was chosen as the preferred substrate for characterization of HEL12 and HEL37. However, this limited the comparisons that were possible within the

known database of AE enzymes. Secondary compounds where analysed to support the hydrolysis of longer carbon chain ester group substitutions within the backbone of the xylan polymer. HEL12 showed that it hydrolyses a butyrate group from the substrate and HEL37 indicated activity when a palmitate group was attached to 4-nitrophenol. This dual function hydrolysis indicates that the enzymes are not only functional on a short ester group, but can also accommodate longer groups. With the secondary substrates used for biochemical analysis varying amongst publicised classification of AE enzymes, one cannot make an accurate comparison between HEL12 and HEL37 with regard to these other substrates, nor relate them to other known AE enzymes. It is however an indication that HEL37, though it is shown to hydrolyse p-nitrophenyl palmitate (C16), was slow at doing so (V_{max} = 0.436 mMol/min/mg protein) compared to p-nitrophenyl acetate (V_{max} = 2.462 mMol/min/mg protein). This compares to the rate that the other AE enzymes break down p-nitrophenyl acetate, which, has a comparatively shorter carbon chain to hydrolyse (C3). It is plausible that the versatility demonstrated by HEL37 has developed due to the harsh environment, and the competition for the energy source present. Thus, regardless of the speed of hydrolysis this makes HEL37 of high interest to industrial hemicellulose hydrolysis.

HEL12 and HEL37 hydrolyses p-nitrophenyl acetate than other documented AE enzymes and therefore a revised protocol had to be designed for kinetic analysis (Section 2.8.2). It also allowed the kinetic assays to be accurate, yet, the substrate concentrations would have to be high to allow surplus substrate. The high substrate concentrations required was at the point where saturation would cause optical interference in the assay. The fast reaction and velocity of enzymes also provided for a quick initial discolouration and reading, and therefore a normalisation of data was pertinent to the study. When HEL12 and HEL37 was stored in 50 mM potassium phosphate and 10% glycerol at 4°C and room temperature (25 ± 1.0°C) after 4 days, 90.79% and 97.71% of activity of the respective enzymes were obtained. This indicates that HEL37 has a higher specific potential as an industrial hydrolytic enzyme than HEL12.

4.5 Native Substrate analysis

The ability of an esterase to hydrolyse short-chain synthetic compounds only hints at the association of the enzyme of being capable of removing the same, or similar subgroups from longer polymeric chains. Due to the sample enrichment during incubation focussing on cellulosic degradation specifically, of which a high percentage (+-40%) is hemicellulose and xylan related, the enzymes were assayed on a variety of plant polymeric substrates [viz. (1) wheat arabinoxylan (Megazyme), (2) xylan from Oats (Sigma-Aldrich), (3) arabinan from sugar beet (Megazyme), (4) xyloglucan from tamarind seed (Megazyme), (5) sucrose acetate isobutyrate (Sigma-Aldrich), (6) chitin from shrimp (Sigma-Aldrich), (7) xylan from beechwood (Sigma-Aldrich)]. Results from this assay show that both enzymes (HEL12 and HEL37) have strong hydrolytic activity for polymeric substrates composing of a xylose homopolymer (samples 2 and 7 in Figures 3.25) as opposed to heteropolymer substrates and other homopolymeric substrates. When HEL37 is compared to HEL12 it is clear that the HEL12 is a more efficient enzyme concerning acetic acid release from polysaccharides and monomeric sucrose acetate isobutyrate. This could be due to the higher velocity rates that was found when applying HEL12 to synthetic compound p-nitrophenyl acetate in the previous assay. The indication that these enzymes both hydrolyse acetyl ester groups from homopolymeric xylan specifically, but not exclusively leads to the conclusion that both enzymes are ambiguous acetyl xylan esterases, capable of hydrolysing an array of polymeric compounds. Comparing this result to the conserved domains, the binding domain of each of these enzymes was clearly seen to be directed to acyl groups. The indels in each of these conserved domains could account for the selective binding to specific conformations to which this acetyl group is attached, which would in this case favour a five carbon sugar, xylose polymer.

Unlike HEL12, HEL37 has no indicated activity on sucrose acetate isobutyrate, though it does however indicate activity on arabinoxylan 0.059 and chitin 0.095. It can be postulated that HEL37 has no binding affinity for a glucose molecule coupled to fructose (sucrose), yet is able, even though only slightly, to hydrolyse the subgroups of homopolymeric glucose substrates such as chitin or heteropolymeric substrate arabinoxylan. This leads to the conclusion that HEL37 might have developed this trait due to some cellulose (glucan) constructs having a low level of acetylation, whereas in plant physiology fructose is rarely used as a predominant monosaccharide in cell wall polymerization, would not necessarily be of high importance to have affinity for.

Acetyl xylan esterases have been characterized from various branches of bacteria for e.g. *Bacillus, Cronobacter, Thermatoga* and *Acinetobacter* species. Thies is the first report of an enzyme isolated from *Citrobacter farmeri* with the ability to hydrolyse acetylated xylan bonds. This merits further study on the characterisation of the hypothetical protein sequence AF41_01806 from *Citrobacter spp MGH55* as an AE enzyme and whether this bacterium is regularly found in degraded industrial saw dust samples.

It was seen that although *B. vallismortis* and *C. farmeri* are from distant phyla of the bacterial kingdom, they are found in the same niche, with similar enzymes that hydrolyse ester bonds on polymers to facilitate increased hydrolysis of the polymer into monomeric subunits as an energy source. Upon comparison of HEL12 and HEL37 to three enzymes in literature: *B. pumilus* PS213, *B. pumilus* AE and *T. maritima* TM0077, several similarities were found in the properties and functions of these enzymes (Levisson *et al*., 2012; Degrassi *et al.,* 1998; Degrassi *et al.,* 2000). Although a pH of 8.0 suits industrial applications, the largest shortcoming of HEL12 and of HEL37 was a lower optimum temperature as their intended industrial application would require. When the enzymes were compared to one another, it was found that they both act in a similar manner and although favouring xylan (hemicellulose), can be utilised to hydrolyse acetyl subgroups from various naturally occurring plant polysaccharides.

5. Conclusion

Plant biomass breakdown into simple sugars for bio-ethanol fermentation has been a goal for many years (Catchpole and Wheeler, 1992). First generation biofuel production has focussed on simple sugar and starch degradation, causing disputes with various humanitarian organisations (Naik *et al.,* 2010; Antizar-Ladislao and Turron-Gomez 2008; Schenk *et al.,* 2008). Second generation biofuel production has focussed on plant cell wall material, cellulose, hemicellulose and lignin to a certain degree, all of which are usually regarded as waste with the exception of fibres utilized for clothing and paper production (Bahcegul *et al.,* 2012; Kapanigowda *et al.,* 2010; Prasifka *et al.,* 2010).

Hemicellulose has been proven difficult to break down, due to side groups attached to the xylose backbone (Biely *et al.,* 2014; Badal, 2003). These side groups are often linked to the backbone by various ester groups (Badal, 2003). Acetyl xylan esterases facilitate hydrolysis of these side groups and are therefore necessary for complete hydrolysis of the hemicellulose into simple sugars.

By improving on acetyl xylan esterase enzyme knowledge and discovery of industrially important acetyl xylan esterases, this study showes that HEL12 and HEL37 have unique properties that are not present in other AE enzymes. The proteins were related to other enzymes and shown to have an elevated V_{max} and a pH in comparison to other AE enzymes characterized on synthetic substrate p-nitrophenyl acetate. The proteins have shown to be genetically similar to other hydrolases, having similar binding groups and have certain deleted and inserted regions that differ from other acetyl xylan esterases.

An important limitation to the use of p-nitrophenyl acetate as characterization substrate is that it is not stable at room temperature, and can infer error within screens that have been performed at room temperature. The substrate is simpler to assay than acetylated xylan, since the length and therefore molar mass of the polymer is unknown, but also does not show the binding and release of acetyl groups from a complex polymer. Further studies on these enzymes included an activity assay on acetylated xylan and other related polymers. Future studies will incorporate cell wall composition of the transgenic *Nicotiana tabacum* created during this study.

References

Abramson M, Shoseyov O, Shani Z (2010) Plant cell wall reconstruction towards improved lignocellulosic production and processability. Plant Science **178**(2):61-72

Ademark P, Varga A, Medev J, Harjunpää V, Drakenberg T, Tjerneld F (1998) Softwood hemicellulose-degrading enzymes from *Aspergillus niger*: Purification of β-mannase. Journal of Biotechnology **63**(3):199-210

Alcaraz LD, Moreno-Hagelsieb G, Eguiarte LE, Souza V, Herrera-Estrella L, Olmedo G (2010) Understanding the evolutionary relationships and major traits of *Bacillus* through comparative genomics. BMC genomics **11**(1):332

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. Journal of Molecular Biology **215:** 403-410

Antizar-Ladislao B, Turrion-Gomez JL (2008) Second-generation biofuels and local bioenergy systems. Biofuel Products and Biorefining **2**:455-469

Antolin G, Tinaut FV, Briceno Y, Castano V, Perez C, Ramırez AI (2002) Optimisation of biodiesel production by sunflower oil transesterification. Bioresource Technology **83**(2):111-114

Arunachalam M, Mohan N, Mahadevan A (2003) Cloning of *Acinetobacter calcoaceticus* chromosomal region involved in catechin degradation. Microbiological Research **158**(1):37-46

Badal CS (2003) Hemicellulose bioconversion. Journal of Industrial Microbiology and Biotechnology **30**(5):279-291

Bahcegul E, Toraman HE, Ozkan N, Bakir U (2012) Evaluation of alkaline pretreatment temperature on a multi-product basis for the co-production of glucose and hemicellulose based films from lignocellulosic biomass. Bioresources Technology **103**:440-445

Baker GC, Smith JJ, Cowan DA (2003) Review and re-analysis of domain-specific 16S primers. Journal of Microbiological Methods **55**(3):541-555

Banerjee G, Car S, Scott-Craig JS, Borrusch MS, Bongers M, Walton JD (2010b) Synthetic multicomponent enzyme mixtures for deconstruction of lignocellulosic biomass. Bioresources Technology **101**:9097-105

Banerjee G, Car S, Scott-Craig JS, Borrusch M, Aslam N, Walton JD (2010c) Synthetic enzyme mixtures for biomass deconstruction: production and optimization of a core set. Biotechnology and Bioengeneering **106**:707-20

Banerjee G, Scott-Craig JS, Walton JD (2010a) Improving enzymes for biomass conversion: a basic research perspective. Bioenergy Resources **3**:82-92

Barbier E (2002). Geothermal energy technology and current status: an overview. Renewable and Sustainable Energy Reviews **6**(1):3-65.

Bartsch S, Kourist R, Bornscheuer UT (2008) Complete inversion of enantioselectivity towards acetylated tertiary alcohols by a double mutant of a *Bacillus subtilis* esterase. Angewandte Chemie International Edition **47**(8):1508-1511

Bauer WD, Talmadge KW, Keegstra K, Alperseim P (1973) The hemicellulose of the walls of suspension-cultured sycamore cells. Plant Physiology **51**(1):174-187

Bencúr P, Steinkellner H, Svoboda B, Muncha J, Strass R, Kolarich D, Hann S, Köllensperger, Glössl J, Altmann F, March L (2005) *Arabidopsis thaliana* β1,2-xylosyltransferase: an unusual glycosyltransferase with the potential to act at multiple stages of the plant *N*-glycosylation pathway. Biochemical Journal **388**(pt2):515-525

Benoit I, Coutinho PM, Schols HA, Gerlach JP, Henrissat B, de Vries RP (2012) Degradation of different pectins by fungi: Correlations and contrasts between the pectinolytic enzyme sets identified in genomes and the growth on pectins of different origins. BMC Genomics **13**:321

Berchmans HJ, Hirata S (2008) Biodiesel production from crude *Jatropha curcas* L. seed oil with a high content of free fatty acids. Bioresource Technology **99**(6):1716-1721

Biely E, Puls J, Schneider H (1991) Acetyl xylan esterases in fungal cellulolytic system. FEBS Letters **186**:80-84

Biely P, Cziszárová M, Agger JW, Li X-L, Puchart V, Vršanská M, Eijsink VGH, Westereng B (2014) *Trichoderma reesei* CE16 acetyl xylan esterase and its role in enzymatic degradation of acetylated hemicellulose. Biochimica et Biophysica Acta **1840**:516-525

Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. Annual Review of Plant Biology **54**:519-546

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical Biochemistry **72**:248-254

Brannon-Peppas L, Blanchette JO (2012) Nanoparticle and targeted systems for cancer therapy. Advanced Drug Delivery Reviews **64**:206-212

Brenner D J, Grimont PA, Steigerwalt AG, Fanning, GR, Ageron E, Riddle CF (1993) Classification of citrobacteria by DNA hybridization: designation of *Citrobacter farmeri* sp. nov., *Citrobacter youngae* sp. nov., *Citrobacter braakii* sp. nov., *Citrobacter werkmanii* sp. nov., *Citrobacter sedlakii* sp. nov., and three unnamed *Citrobacter genomospecies*. International Journal of Systematic Bacteriology **43**(4):645-658

Breton C, Mucha J, Jeanneau C (2001) Structural and functional features of glycosyltransferases. Biochimie **83**(8):713-718

Blum D, Li X, Chen H, Ljungdahl LG (1999) Characterization of an Acetyl Xylan esterase from the Anaerobic Fungus *Orpinomyces* sp. Strain PC-2. Applied and Environmental Microbiology **65**(9):3990-3995

Brown RM jr. (1999) Cellulose structure and biosynthesis. Pure Applied Chemistry **71**(5):767-775

Bugg TDH, Ahmad M, Hardiman EM, Sing R (2011) The emerging role of bacteria in lignin degradation and bio-product formation. Current Opinion in Biotechnology **22**:394-400

Caffall KH, Mohnen D (2009) The structure, function, and biosynthesis of plant cell wall pectin polysaccharides. Carbohydrate Research **344**(14):1879-1900

Casaburi A, Villani F, Toldrá F, Sanz,Y (2006) Protease and esterase activity of *staphylococci*. International Journal of Food Microbiology **112**(3):223-229.

Catchpole WR, Wheeler CJ (1992) Estimating plant biomass: a review of techniques. Australian Journal of Ecology **17**(2):121-131

Chen J, Liang R-H, Lui W, Lui C-M, Li T, Tu Z-C, Wan J (2012) Degradation of high methoxyl pectin by dynamic high pressure microscopy and its mechanism. Food Hydrocolloids **28**:121-129

Cherif S, Mnif S, Hadrich F, Abdelkafi S, Sayadi S (2011) A newly high alkaline lipase: an ideal choice for application in detergent formulations. Lipids in Health and Disease **10**(1):1-8.

Christgau S, Sandal T, Kofod LV, Dalbøge H (1995) Expression cloning, purification and characterization of a β-1, 4-glactanase from *Aspergillus aculeatus*. Current Genetics **27**(2):135-141

Clay J (2003) World agriculture and the environment: a commodity-by-commodity guide to impacts and practices. Washington, DC [etc.]: Island

Creissen G, Reynolds H, Xue Y, Mullineaux P (1995) Simultaneous targeting of pea glutathione reductase and of a bacterial fusion protein to chloroplasts and mitochondria in transgenic tobacco. The Plant Journal **8**(2):197-175

Dartois V, Baulard A, Schanck K, Colson C (1992) Cloning, nucleotide sequence and expression in *Escherichia coli* of a lipase gene from *Bacillus subtilis* 168. Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression **1131**(3):253-260.

de Graaff L, van den Broeck H, Visser J (1992). Isolation and characterization of *the Aspergillus niger* pyruvate kinase gene. Current genetics **22**(1):21-27.

De Vries RP, Visser J (2001) *Aspergillus* enzymes involved in degradation of plant cell wall polysaccharides. Microbiology Molecular Biology Review **65**(4):497-522

De Vries JA, Rombouts FM, Voragen AGJ, Pilnik W (1982) Enzymatic degradation of apple pectins. Carbohydrate Polymers **2**(1):25-33

Demirbaş A (2006) Global Renewable Energy Resources. Energy Sources **28**(8):779-792

Degrassi G, Kojic M, Ljubijankic G, Venturi V (2000) The acetyl xylan esterase of *Bacillus pumilus* belongs to a family of esterases with broad substrate specificity. Microbiology **146**(7):1585-1591

Degrassi G, Okeke BC, Bruschi C V, Venturi V (1998) Purification and characterization of an acetyl xylan esterase from *Bacillus pumilus*. Applied and Environmental Microbiology, **64**(2):789-792

Drudy D, O'Rourke M, Murphy M, Mullane NR, O'Mahony R, Kelly L, Fisher M, Sanjaq S, Shannon P, Wall P, O'Mahony M, Whyte P, Fanning S (2006) Characterization of a collection of *Enterobacter sakazakii* isolates from environmental and food sources. International Journal of Food Microbiology, **110**(2):127-134.

Du L, Huo Y, Ge F, Yu J, Li W, Cheng G, Yong B, Zeng L, Huang M (2010) Purification and characterization of novel extracellular cholesterol esterase from *Acinetobacter* sp. Journal of basic microbiology **50**(S1):S30-S36.

Dutta S and Ray L (2009) Production and characterization of an alkaline thermostable crude lipase from an isolated strain of *Bacillus cereus C7*. Applied Biochemistry and Biotechnology **159**(1):142-154

Eggert T, Pencreac'h G, Douchet I, Verger R, Jaeger K E (2000) A novel extracellular esterase from *Bacillus subtilis* and its conversion to a monoacylglycerol hydrolase. European Journal of Biochemistry **267**(21):6459-6469

Elleuch M, Bedigian D, Roiseux O, Besbes S, Blecker C, Attia H (2011) Dietary fibre and fibre-rich by-products of food processing: characterisation, technological functionality and commercial applications: A review. Food Chemistry **124**:411-421

Emanuelsson O, Brunak S, von Heijne G, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP and related tools. Nature Protocols **2**:953-971

Emanuelsson O, Nielsen J, Brunak S, von Heijne G (2000) Predicting subcellular localization of proteins Based on their N-terminal amino acid sequence. Journal of Molecular Biology **300**(4):1005-1016

Faye L, Gomord V, Fitchettelaine AC, Chrispeels MJ (1993) Affinity purification of antibodies specific for Asn-linked glycans containing α1→3 fucose or β1→2 xylose. Analytical biochemistry **209**(1):104-108

Ferreira LM, Wood T M, Williamson G, Faulds C, Hazlewood GP, Black GW, Gilbert H J (1993) A modular esterase from *Pseudomonas fluorescens* subsp. cellulose contains a non-catalytic cellulose-binding domain. Biochemistry Journal **294:**349-355

Fontana L, Atella V, Kammen DM (2013) Energy efficiency as a unifying principle for human, environmental, and global health. F1000 Research doi: 10.12688/f1000research.2-101.v1

Fraiture C, Giordano M, Liao Y (2007) Biofuels and implications for agricultural water use: blue impacts and green energy. Water Policy **10**(1):67-81

Gordon CAAH, Lomax JA (1983) Substituent groups linked by alkali-labile bonds to arabinose and xylose residues of legume, grass and cereal straw cell walls and their fate during digestion by rumen microorganisms. Journal of the Science of Food and Agriculture **34**:1330-1340

Grauers A (1996) Efficiency of three wind energy generator systems. Energy Conversion, IEEE Transactions on, **11**(3):650-657

Grimm M, Stephan R, Iversen C, Manzardo GG, Rattei T, Riedel K, Reupp A, Frishman D, Lehner A (2008) Cellulose as an extracellular matrix component present in *Enterobacter sakazakii* biofilms. Journal of Food Protection® **71**(1):13-18

Hall, TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series **41**:95–98

Han SJ, Yoo YJ, Kang HS (1995) Characterization of a bifunctional cellulase and its structural gene. The cell gene of *Bacillus* sp. D04 has exo- and endoglucanase activity. Journal of Biological Chemistry **270**(43):26012-26019

Hasan F, Shah AA, Hameed A (2006) Industrial applications of microbial lipases. Enzyme and Microbial Technology **39**(2):235-251

Havlik P, Schneider, UA, Schmid E, Böttcher H, Fritz S, Skalský R, Aoki K, De Cara S, Kindermann G, Kraxner F, Leduc S, McCallum I, Mosnier A, Sauer T, Obersteiner M (2011) Global land-use implications of first and second generation biofuel targets. Energy Policy **39**(10):5690-5702

Henrissat B, Bairoch A (1993) New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. Biochemistry Journal **293:**781–788

Henrissat B, Bairoch A (1996) Updating the sequence-based classification of glycosyl hydrolases. Biochemistry Journal **316**:695-696

Herrera-Estrella A, Van Montagu M, Wang K (1990) A bacterial peptide acting as a plant nuclear targeting signal: the amino-terminal portion of Agrobacterium VirD2 protein directs a beta-galactosidase fusion protein into tobacco nuclei. PNAS **87**(24): 9534-9537

Holmes S (2003) Bootstraping phylogenetic trees: theory and methods. Statistical science **18**(2):241-255

Igarashi K, Uchihashi T, Koivula A, Wada M, Kimura S, Okamoto T, Penttilä M, Ando T, Samehima M (2011) Traffic jams reduce hydrolytic efficiency of cellulases on cellulose surface. Science **6047**(330):1279-1282

Irwin DC, Spezio M, Walker LP, Wilson DB (1993) Activity studies of eight purified cellulases: specificity, synergism, and binding domain effects. Biotechnology and Bioengeneering **42**(8):1002-1013

Iversen C, Mullane N, McCardell B, Tall BD, Lehner A, Fanning S, Stephan, R Joosten H (2008) *Cronobacter* gen. nov., a new genus to accommodate the biogroups of *Enterobacter sakazakii*, and proposal of *Cronobacter sakazakii* gen. nov., comb. nov., *Cronobacter malonaticus* sp. nov., *Cronobacter turicensis* sp. nov., *Cronobacter muytjensii* sp. nov., *Cronobacter dublinensis* sp. nov., *Cronobacter genomospecies* 1, and of three subspecies, *Cronobacter dublinensis* subsp*. dublinensis* subsp. nov., *Cronobacter dublinensis* subsp. *lausannensis* subsp. nov. and *Cronobacter dublinensis subsp. lactaridi* subsp. nov. International Journal of Systematic and Evolutionary Microbiology **58**(6):1442-1447.

Jandas PJ, Mohanty S, Nayak SK (2013) Surface treated banana fiber reinforced poly (lactic acid) nanocomposites for disposable applications. Journal of Cleaner Production **52**:392-401

Joseph S, Forsythe SJ (2011) Predominance of *Cronobacter sakazakii* sequence type 4 in neonatal infections. Emerging Infectious Diseases, **17**(9):1713.

Kapanigowda M, Stewart BA, Howell TA, Kadasrivenkata H, Baumhardt RL (2010) Growing maize in clumps as a strategy for marginal climatic conditions. Field Crop Research **188**:115-125

Keegstra K (2010) Plant cell walls. Plant Physiology, **154**(2):483-486

Khan AW, Lamb KA, Overend RP (1990) Comparison of natural hemicellulose and chemically acetylated xylan as substrates for determination of acetyl-xylan esterase activity in *Aspergilli*. Enzyme Microbiology and Technology **12**:127-130

Khunrong T, Punsuvon V, Vaithanomsat P, Pomchaitaward C (2011) Production of ethanol from pulp obtained by steam explosion pretreatment of oil palm trunk. Energy Sources Part A 33:221-228.

Kim JY, Kavas M, Fouad WM, Nong G, Preston JF, Altpeter F (2010) Production of hyperthermostable GH10 xylanase Xyl10B from *Thermotoga maritima* in transplastomic plants enables complete hydrolysis of methylglucuronoxylan to fermentable sugars for biofuel production. Plant Molecular Biology 2010. doi:10.1007/s11103-010-9712-6.

Kim S, Dale B (2006) Ethannol fuels: E10 or E85 – Life Cycle Perspectives. The International Journal of Life Cycle Assessment **11**(2):117-121

Kirk JL, Beaudette LA, Hart M, Moutoglis P, Klironomos JN, Lee H, Trevors JT (2004) Methods for studying soil microbial diversity. Journal of Microbial Methods **58**:169-188

Kjer-Neilsen L, van Vliet C, Erlich R, Toh BH, Gleeson PA (1999) The Golgi-targeting sequence of the peripheral membrane protein p230. Journal of Cell Science **112**:1645-1654

Kleene R, Berger EG (1993) The molecular and cell biology of glycosyltransferases. Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes **1154**(3):283-325.

Klemm D, Heublein B, Fink HP, Bohn A (2005) Cellulose: fascinating biopolymer and sustainable raw material. Angewandte Chemie International Edition **44**(22):3358-3393

Kossmann J, Lorberth R (2010) Plants which synthesize a modified starch process for the production thereof and modified starch. US Patent US 2010/0174061 A1

Kosugi A, Morasima K, Doi RH (2002) Xylanase and AEtyl xylan esterase activities of XynA, a key subunit of the *Clostridium cellulovorans* cellulosome for xylan degradation. Applied and Environmental Microbiology **68**(12):6399-6402

Kouker G, Jaeger KE (1987) Specific and sensitive plate assay for bacterial lipases. Applied and Environmental Microbiology **53**(1):211-213

Kleber-Janke T, Becker W-M (2000) Use of modified BL21 (DE3) *Eschericia coli* cells for high-Level expression of recombinant peanut allergens affected by poor codon usage. Elsevier **3**(19):419-424

Kumar R, Singh S, Singh OV (2008) Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives. Journal of Indian Microbiology and Biotechnology **35**:377-391

Lane DJ, Pace B, Olsen GJ, Stahl DA, Sogin ML, Pace NR (1985) Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. Proceedings of the National Academy of Sciences **82**(20):6955-6959

Lau MW, Gunawan C, Dale BE (2009) The impacts of pretreatment on the fermentability of pretreated lignocellulosic biomass: a comparative evaluation between ammonia fiber expansion and dilute acid pretreatment. Biotechnology Biofuels **2**:30

Lerouge P, Cabanes-Macheteau M, Rayon C, Fischette-Lainé AC, Gomord V, Faye L (1998) N-glycoprotein biosynthesis in plants: recent developments and future trends. Plant Molecular Biology **38**(1-2):31-48

Levisson M, Han GW, Deller MC, Xu Q, Biely P, Hendriks S, Eyck LFT, Felsenburg C, Roversi P, Miller MD, McMullan D, von Delft F, Kreusch A, Deacon AM, van der Oost J, Lesley SA, Elsliger M-A, Kengen SWM, Wilson IA (2012) Functional and structural characterization of thermostable acetyl esterase from *Thermotoga maritima.* Proteins **80**:1545-1559

Liu S (2010) Woody biomass: niche position as a source of sustainable renewable chemicals and energy and kinetics of hot-water extraction/hydrolysis. Biotechnology Advances **28**:563–82

Liu X, Ding S (2009) Molecular characterization of a new acetyl xylan esterase (AEII) from edible straw mushroom *Volvariella volvacea* with both de-O-acetylation and de-N-acetylation activity. FEMS Microbiology Letters **295**:50-56

Lund JW Freeson DH (2001) World-Wide directed uses of geothermal energy 2000. Geothermics **30**(1):29-68

Lüthi E, Love DR, McAnulty J, Wallace C, Caughey PA, Saul D, Bergquist PL (1990). Cloning, sequence analysis, and expression of genes encoding xylan-degrading enzymes from the thermophile" *Caldocellum saccharolyticum*". Applied and environmental microbiology **56**(4):1017-1024

Lv W, Yu Z (2012) Isolation and characterization of two thermophilic cellulytic strains *Clostridium thermocellum* from a compost sample. Journal of Applied Microbiology **114**:1001-1007

Ma J, Zhang Z, Wang B, Kong X, Wang Y, Cao S, Feng Y (2006) Overexpression and characterization of a lipase from *Bacillus subtilis*. Protein expression and Purification **45**(1):22-29

Margolles-Clark E, Jacques-Silva MC, Ganesan L, Umland O, Kenyon NS, Ricordi C, Berggern P, Buchwald P (2009) Suramin inhibits the CD40–CD154 stimulatory interaction: a possible mechanism for immunosuppressive effects. Biochemical Pharmacology **77**(7):1236-1245

Mathew GM, Lin SJ, Chang JJ, Huang CC (2011). DGGE detection and screening of lignocellulolytic bacteria from the termite gut of *Coptotermes formosanus*. Malays Journal of Microbiology **7**:201-209

Merino ST, Cherry J (2007) Progress and challenges in enzyme development for biomass utilization. Advanced Biochemical Engineering and Biotechnology **108**:95-120

Mohnen D (2008) Pectin structure and biosynthesis. Current Opinion in Plant Biology **11**:266-277

Mok WSL, Antal MJ Jnr (1992) Uncatalyzed solvolysis of whole biomass hemicellulose by hot compressed liquid water. Industrial and Engineering Chemistry Research **31**(4):1115-1161

Mu Y, Xiu ZL, Zhang DJ (2008) A combined bioprocess of biodiesel production by lipase with microbial production of 1,3-propanediol by *Klebsiella pneumoniae*. Biochemical Engineering Journal 40(3):537-541

Nakamura A, Furura H, Maeda H, Takao T, Nagamatsu Y (2002) Analysis of the molecular construction of xylogalacturononan isolated from soluble soybean polysaccharides. Bioscience, Biotechnology and Biochemistry **66**(5):1155-1158

Naik SN, Goud VV, Rout PK, Dalai AK (2010) Production of first and second generation biofuels: a comprehensive review. Renewable and Sustainable Energy Reviews, **14**(2):578-597

Nair IC, Jayachandran K, Shashidhar S (2007) Treatment of paper factory effluent using a phenol degrading *Alcaligenes* sp. under free and immobilized conditions. Bioresource Technology **98**(3):714-716.

Nilsson IM, Von Heijne G (1993) Determination of the distance between the oligosaccharyltransferase active site and the endoplasmic reticulum membrane. Journal of Biological Chemistry **268**(8):5798-5801

Nishiyama Y, Langan P, Chanzy H (2002) Chrystal structure and hydrogen-bonding system in Cellulose 1B from synchroton X-ray and Neutron Fiber Diffraction. Journal of American Chemical Society **124**(31):9074-9082

Oberoi HS, Rawat R, Chadha BS (2014) Response surface optimization for enhanced production of cellulases with improved functional characteristics by new isolated *Aspergillus niger* HN-2. Antonie van Leeuwenhoek **105**:119-135

Oetting WS, Lee HK, Flanders DJ, Wisner GL, Sellers TA, King RA (1995) Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed primers. Genomics **30**:450-458

Okolie PN, Ugochukwa EN (1988) Changes in activities of cell wall degrading enzymes during fermentation of cassava (*Manihot esculenta* Crantz) with *Citrobacter freundii*. Journal of the Science of Food and Agriculture **44**(1):51-61

O'Sullivan AC (1996) Cellulose: the structure slowly unravels. Cellulose **4**:173-207

Pagny S, Bouissonnie F, Sarkar M, Follet-Gueye ML, Driouich A, Schachter H, Faye L, Gomord V (2003) Structural requirements for *Arabidopsis* β1, 2-xylosyltransferase activity and targeting to the Golgi. The Plant Journal **33**(1):189-203

Park S, Baker JO, Himmel ME, Parilla PA, Johnson DK (2010) Cellulose crystallinity index: measurement techniques and their impact in terpreting cellulase performance. Biofuels **3**:1-10

Pauly M, Keegstra K (2010) Plant cell wall polymers and precursors for biofuels. Current Opinion in Plant Biotechnology **13**(3):304-311

Peng X-w, Ten J-l, Zhong L-x, Sun R-c (2011) Nanocomposite films based on xylan-rich hemicelluloses and cellulose nanofibers with enhanced mechanical properties. Bio Macromolecules **12**:3321-3329

Pérez-Lombard L, Oritz J, Pout C (2008) A review on buildings energy consumption information. Energy and Buildings **40**(3):394-398

Persson M, Bornscheuer UT (2003) Increased stability of an esterase from *Bacillus stearothermophilus* in ionic liquids as compared to organic solvents. Journal of Molecular Catalysis B: Enzymatic **22**(1):21-27

Podschun R Ullmann U (1998) *Klebsiella* spp. as nosocomial pathogens: epidemiology, taxonomy, typing methods, and pathogenicity factors. Clinical microbiology reviews **11**(4):589-603

Popper ZA, Fry SC (2005) Widespread occurrence of a covalent linkage between xyloglucan and acidic polysaccharides in suspension-cultured angiosperm cells. Annuals of Botany **96**:91-99

Posada D (2008) jModelTest: Phylogenetic Model Averaging. Molecular biology and Evolution **25**(7):1253–1256

Prasifka JR, Bradsaw JD, Boe AA, Lee D, Adamski D, Grey ME (2010) Symptoms, Distribution and abundance of stem-boring caterpillar, *Blastobasis repartella* (Dietz), in Switchgrass. Bioengeneering Resources **3**:238-242

Ragaglini G, Triana F, Villani R, Bonari E (2011) Can sunflower provide biofuel for inland demand? An integrated assessment of sustainability at regional scale. Energy, **36**(4):2111-2118

Raikhel N (1992) Nuclear Targeting in Plants**.** Plant Physiology **100**(4):1627-1632

van Ree R, Cabanes-Macheteau M, Akkerdaas J, Milazzo JP, Loutelier-Bourhis C, Rayon C, Villalba M, Koppelman S, Aalberse R, Rodriguez R, Lerouge, P (2000) β (1, 2)-xylose and α (1, 3)-fucose residues have a strong contribution in IgE binding to plant glycoallergens. Journal of Biological Chemistry **275**(15):11451-11458

Reijnders L (2010) The life cycle emission of greenhouse gasses associated with plant oils used as Biofuel. Renewable Energy **36**:879-880

Ridley BL, O'Neill MA, Mohnen D (2001) Pectins: structure, biosynthesis and oligogalacuronide-related signalling. Pytochemistry **57**:929-967

Riley M, Abe T, Arnaud, M B, Berlyn M K, Blattner F R, Chaudhuri R R, Glasner JD, Horiuchi T, Keseler IM, Kosuge T, Mori H, Perna NT, Plunkett G, Rudd KE, Serres MH, Thomas GH, Thomason NR, Wishart D, Wanner B L (2006) *Escherichia coli* K-12: a cooperatively developed annotation snapshot—2005. Nucleic Acids Research **34**(1):1-9

Roberts MS, Nakamura LK, Cohan FM (1996) Bacillus vallismortis sp. nov., a close relative of *Bacillus subtilis*, isolated from soil in Death Valley, California. International journal of systematic Bacteriology **46**(2):470-475

Rollin JA, Zhu Z, Sathitsuksanoh N, Zhang YHP (2011) Increasing cellulose accessibility is more important than removing lignin: a comparison of cellulose solvent-based lignocellulose fractionation and soaking in aqueous ammonia. Biotechnology and Bioengeneering **108**:22–30

Rosgaard L, Pedersen S, Langston J, Akerhielm D, Cherry JR, Meyer AS (2007) Evaluation of minimal *Trichoderma reesei* cellulase mixtures on differently pretreated barley straw substrates. Biotechnology Progress **23**:1270–6

Saha BC (2003) Hemicellulose bioconversion. Journal of Industrial Microbiology and Biotechnology **30**(5):279-291

Saitou N, Nei M (1987) The neighbour-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution **4**(4):406-425

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: a laboratory manual. 2nd ed. N.Y., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, 1989. 1659 p. ISBN 0-87969-309-6.

Sanderson K (2011) Lignocellulose: A chewy problem. Nature **474**:S12-S14

Schenk PM, Tomas-Hall SR, Stephens E, Marx UC, Mussgnug JH, Posten C, Kruse O, Hanekamer B (2008) Second generation biofuels: high efficiency microalgae for biodiesel production. Bioenergy Research **1**:20-43

Schimpf U, Hanreich A, Mähnert P, Unmack T, Junne S, Renpenning J, and Lopez-Ulibarri R (2013) Improving the efficiency of large scale biogas processes: pectinolytic enzymes accelerate the lignocellulose degradation. Journal of Sustainable Energy and Environment **4**:53-60

Shabeb MSA, Younis MAM, Hezayen FF, Nour-Eldein MA (2010) Production of cellulase in low-cost medium by *Bacillus subtilis* KO strain. World applied sciences journal **8**(1):35-42

Shafiee S, Topal E (2009) When will fossil fuel reserves be diminished? Energy Policy **37**(1):181-189

Shah P, Gutierrez-Sanchez G, Orlando R, Bergmann C (2009) A proteomic study of pectin-degrading enzymes secreted by *Botrytis cinerea* grown in liquid culture. Proteomics **9**(11):3126-3135

Sjöling S, Glaser E (1998) Mitochondrial targeting peptides in plants. Trends in Plant Science **3**(4):136-140

Sorek N, Yeasts TH, Szemenyei H, Youngs H, Somerville CR (2014) The implications of lignocellulosic biomass chemical composition for the production of advanced biofules. Bioscience Advance Acess doi:10.1093/biosci/bit037

Subhadra BG (2010) Sustainability of algal biofuel production using integrated renewable energy park (IREP) and algal biorefinery approach. Energy Policy **38**:5892-5901

Takahashi T, Satoh I, Kikuchi N (1999) NOTE Phylogenetic relationships of 38 taxa of the genus *Staphylococcus* based on 16S rRNA gene sequence analysis. International Journal of Systematic Bacteriology **49**(2):725-728

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance and Maximum Parsimony. Molecular Biology and Evolution doi:10.1093/molbev/msr121

Thompson DJ, Higgins DG, Gibson TJ (1994) ClustelW: improving the sensitivity of progress multiple sequence alignment through sequence weighting, positioning-specific gap penalties and weight matrix choice. Nucleic Acids Research **22**(22):4673-4680

Tian J, Wang J, Zhao S, Jiang C, Zhang X, Wang X (2009) Hydrolysis of cellulose by the heteropoly acid $H_3PW_{12}O_{40}$. Cellullose **17**:587-594

Toth K, Van Gool MP, Schols HA, Sammuels GJ, Gruppen H, Szakacs G (2013) Diversity in production of xylan-degrading enzymes among species belonging to the *Trichoderma* section *Longibachaitum*. Bioengeneering Resources **6**:631-643

Twidell J, Weir T (2006). *Renewable Energy Resources*. 2nd ed. **New York: Taylor and Francis** 1-487

Van den Broeck G, Timko MP, Kausch AP, Cashmore AT, Van Montagu M, Herra-Estrella L (1985) Targeting of a foreign protein to chloroplasts by fusion to the transit peptide from the small subunit of ribulose 1,5-bisphosphate carboxylase. Nature **313**:358-363

Vanholme R, Demedst B, Morreel K, Ralph J, Boerjan W (2010) Lignin biosynthesis and Structure. Plant Physiology **153**:895-905

Verma D, Kawarabayasi Y, Miyazaki K, Satyanarayana T (2013) Cloning, expression and characteristics of a novel alkalistable and thermostable xylanase encoding gene (Mxyl) retrieved from compost-soil metagenome. PLoS ONE **8**(1): e52459. doi:10.1371/journal.pone.0052459

Visca P, Seifert H, Towner KJ (2011) *Acinetobacte*r infection–an emerging threat to human health. IUBMB life **63**(12):1048-1054

Vitale A, Raikhel NV (1999) What do proteins need to reach different vacuoles? Trends in Plant Science **4**(4):149-155

Wang PY, Tsai SW (2009) Hydrolytic resolution of R-3-hydroxy-3-phenylpropionates by esterase from *Klebsiella oxytoca*: Effects of leaving alcohol, covalent immobilization and aqueous pH. Journal of Molecular Catalysis B: Enzymatic **59**(1):70-75

Waters DM, Murray PG, Miki Y, Martinez AT, Tuohy MG, Faulds CB (2012) Cloning, overexpression in *Escherichia coli*, and characterization of a thermostable fungal acetylxylan Esterase from *Talaromyces emersonii*. Applied and Environmental Microbiology **78**(10):3759-3762

Walker LP, Belair CD, Wilson DB, Irwin DC (1993) Engineering cellulase mixtures by varying themole fraction of *Thermomonospora fusca* E5 and E3, *Trichoderma reesei* CBHI, and *Caldocellum saccharolyticum* β-glucosidase. Biotechnology and Bioengeneering **42**:1019–28

Weinstock G, Sodergren E, Lobos EA, Fulton L, Fulton R, Courtney L, Fronick C, O'Laughlin M, Godfrey J, Wilson RM, Miner T, Farmer C, Delehaunty K, Cordes M, Minx P, Tomlinson C, Chen J, Wollam A, Pepin KH, Bhonagiri V, Zhang X, Warren W, Mitreva M, Mardis ER, Wilson RK (2013) Direct NCBI Submission. Submitted (21-JUN-2013) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park, St. Louis, MO 63108, USA

Willats WGT, Steele-King CG, Marcus SE, Knox JP (1999) Side chains of pectic polysaccharides are regulated in relation to cell proliferation and cell differentiation. The Plant Journal **20**: 619–628.

Xiao ZZ, Zhang X, Gregg DJ, Saddler JN (2004) Effects of sugar inhibition on cellulases and beta-glucosidase during enzymatic hydrolysis of softwood substrates. Applied Biochemistry and Biotechnology **113**:1115–1126

Xiros C, Topakas E, Christakapoulos P (2013) Hydrolysis and fermentation for cellulosic ethanol production. WIREs Energy Environment **2**:633-654

Yakimov mM, Timis KN, Wray V, Fredrickson HL (1995) Characterization of a new lipopeptide surfactant produced by thermotolerant and halotolerant subsurface *Bacillus licheniformis* BAS50. Applied and Environmental Microbiology **61**(5):1706-1713

Yamaguchi T, Dulubova I, Min S, Chen X, Rizo J, Südhof TC (2002) Sly1 Binds to Golgi and ER Syntaxins via a Conserved N-Terminal Peptide Motif. Developmental Cell **3**(2)295-305

Yang H, Yan R, Chen H, Lee DH, Zheng C (2007) Characteristics of hemicellulose, cellulose, lignin pyrolysis. Fuel **86**(12): 1781-1788

Yu EK, Deschatelets L, Louis-Seize G, Saddler JN (1985) Butanediol production from cellulose and hemicellulose by *Klebsiella pneumoniae* grown in sequential co-culture with *Trichoderma harzianum*. Applied and environmental microbiology **50**(4):924-929

Zaldivar J, Nielsen J, Olsson L (2001) Fuel ethanol production from lignocellulose: a challenge for metabolic engineering and process of integration. Applied Microbiology and Biotechnology **56:**17-34

Zhang J, Zhang B, Zhang J, Lin L, Liu S, Ouyang P (2010a) Effect of phosphoric acid pretreatment on enzymatic hydrolysis of microcrystalline cellulose. Biotechnology Advances **28**:613-619.

Zhang M, Su R, Qi W, He Z (2010b) Enhanced enzymatic hydrolysis of lignocellulose by optimizing enzyme complexes. Applied Biochemistry and Biotechnology **160**:1407-1414

Zhang Z, Donaldson AA, Ma X (2012) Advances and future directions in enzyme technology for biomass conversion. Biotechnology Advances **30**:913-919

Znameroski EA, Coradetti ST, Roche CM, Tsai JC, Lavarone AT, Cate JHD, GlassNJ (2012) Introduction of lignocellulose-degrading enzymes in *Neurospora crassa* by cellodextrins. PNSAS **109**:6012-6017

Table A1: Summary of all the primer pairs used in the study

Primer Pair	Name	Forward Primer	Reverse Primer	Use	Reference
1	HEL1	5'-CCCCAGATCTATGATCGAACTGCAAACAGA-3' *Bgl* II	5'-GGGGGTCGACTCAGAGATGACGGCGGAAAAA-3' *Sal* II	Bacterial Expression/Sequencing	
2	HEL12	5'-CCCCAGATCTATGATTGAGATCGACAATAT-3' *Bgl* II	5'-CCCCGTCGACCTAAAGGTGTGTGGTAAAAA-3' *Sal* II	Bacterial Expression/Sequencing	
3	HEL 21	5'-GGGGAGATCTATGATTGCACTAGAAATGCG-3' *Bgl* II	5'-GGGGGTCGACTTAAAGATGCTGCTGAAAA-3' *Sal* II	Bacterial Expression/Sequencing	
4	HEL37	5'-CCCCAGATCTATGGCACTCATGACAATC-3' *Bgl* II	5'-CCCCCGTCGACGGTTAGTAGAAGAATAGG-3' *Sal* II	Bacterial Expression/Sequencing	
5	Xyl-T-36	5'-GGGGGGTACCATGAGTAAACGGAATCCGAA-3' *Kpn* I	5'-CCCCAGATCTCGGTGAAAACGACGATGAGT-3' *Bgl* II	Plant Fusion Protein	
6	Xyl-T-36	5'-GGGGGCGGCCGCATGAGTAAACGGAATCCGAA-3' *Not* I	5'-CCCCGGTA^CCGGTGAAAACGACGATGAGT-3' *Kpn* I	Plant Fusion Protein	
7	HEL12	5'-CCCCGGTACCATGATTGAGATCGACAATAT-3' *Kpn* I	5'-CCCCGTCGACCTAAAGGTGTGTGGTAAAAA-3' *Sal* II	Plant Expression	
8	HEL 37	5'-CCCCGGTACCATGGCACTCATGACAATC-3' *Kpn* I	5'-CCCCCGTCGACGGTTAGTAGAAGAATAGG-3' *Sal* II	Plant Expression	
9	E9F	5'-GAGTTTGATCCTGGCTCAG-3'	5'-ACCGCGGCKGCTGGC-3'	16S Amplification	Stackebrandt and Goebel, 1994
10	F8	5'-CACGGATCCAGACTTTGATKYMTGGCT-3'	5'-GTGAAAGCTTACGGYTAGCTTGTTACG-3'	16S Amplification	Stackebrandt and Goebel, 1995
11	M13	5'-TCACACAGGAAACAGCTATGAC-3'	5'-TCACACAGGAAACAGCTATGAC-3'	Sequencing/Screening Primer	Oetting *et al.,* 1995
12a	CMV-35S	5'-CGCAAATGGGCGGTAGGCGTG-3'		Sequencing/Screening Primer	Matsuoka *et al.,* 2002
12b	T7 Promotor	5'-TAATACGACTCACTATAGGG-3'		Sequencing/Screening Primer	Ogura *et al.,* 1999
12c	T7 Terminater		5'-GCTAGTTATTGCTCAGCGG-3'	Sequencing/Screening Primer	Ogura *et al.,* 1999
13a	OCS-Terminator		5'-CGCAAATGGGCGGTAGGCGTG-3'	Sequencing/Screening Primer	Zhang and Singh, 1994
13b	pJET 1.2 Reverse		5'-AAGAACATCGATTTTCCATGGCAG-3'	Sequencing/Screening Primer	Schartz *et al.,* 2011

Figure A1. cDNA library clone 12 ORF HEL12 with engineered restriction sites added

Figure A2. Translated protein sequence from HEL12 ORF

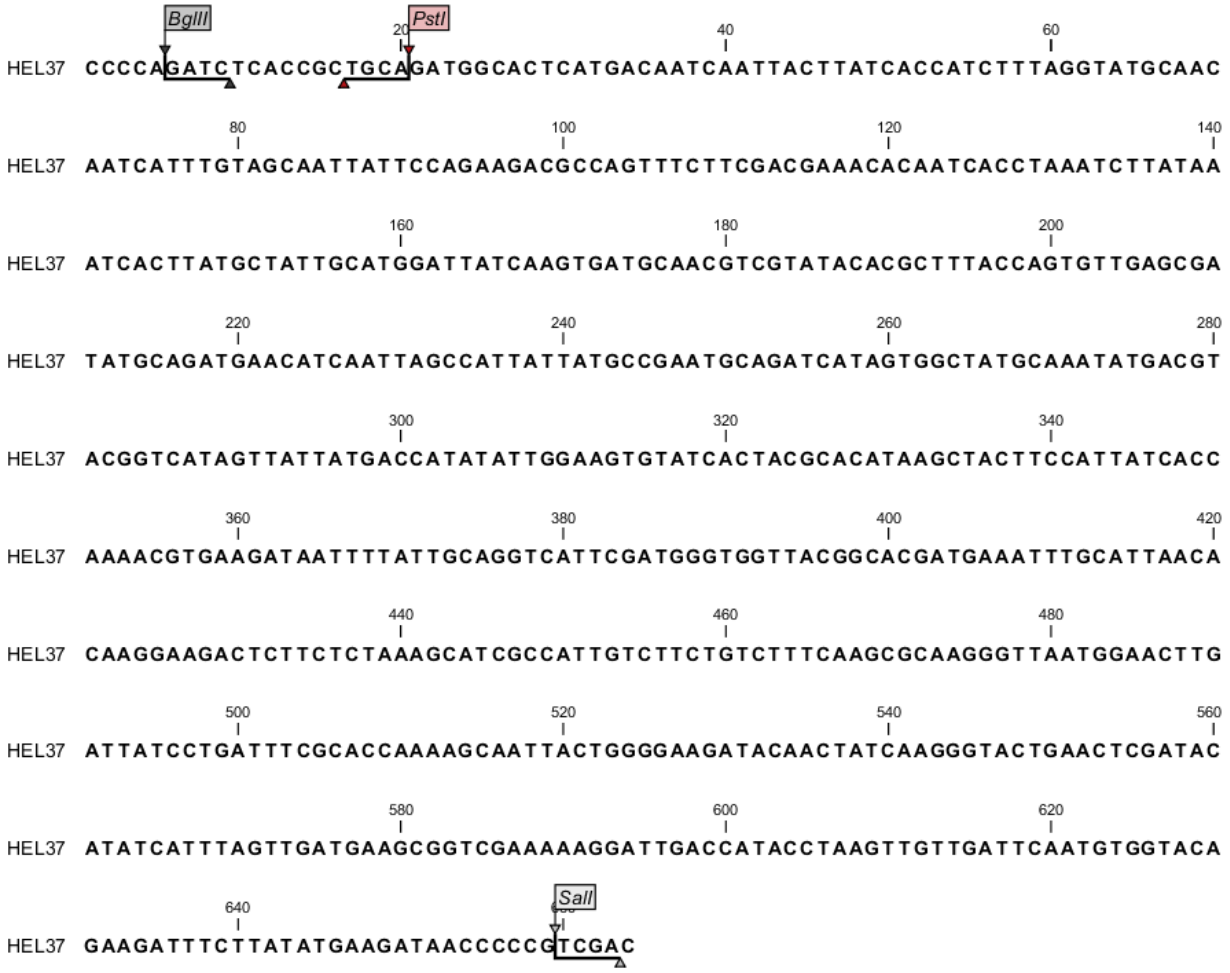

Figure A3. cDNA library clone 37 ORF HEL37 with engineered restriction sites added

Figure A4. Translated protein sequence from HEL37 ORF

Figure A5. Partial sequence of *A.thaliana* β-(1,2)-Xylosyltransferase, Underlined area indicates *Xyl-T36* amplified sequence

Table A2. Table of constructed plasmid constructs

Construct	Vector	Instert	Insert sites	Use of vector
pBK-CMV:Library	pBKCMV	Library fragments	*BamH*I	Activity screening in *E.coli*
pJET:HEL1	pJet	Amplified ORF (HEL1)	T-overhang insert	Cloning and redunancy
pJET:HEL12	pJet	Amplified ORF (HEL12)	T-overhang insert	Cloning and redunancy
pJET:HEL21	pJet	Amplified ORF (HEL21)	T-overhang insert	Cloning and redunancy
pJET:HEL37	pJet	Amplified ORF (HEL37)	T-overhang insert	Cloning and redunancy
pJET:Xyl-T36	pJet	A. *thaliana* Xyl-T Targeting peptide	T-overhang insert	Cloning and redunancy
pGEM-T-Easy:HEL1	pGEM-T-Easy	Amplified ORF (HEL1)	Blunt insert	Cloning
pGEM-T-Easy:HEL12	pGEM-T-Easy	Amplified ORF (HEL12)	Blunt insert	Cloning
pGEM-T-Easy:HEL21	pGEM-T-Easy	Amplified ORF (HEL21)	Blunt insert	Cloning
pGEM-T-Easy:HEL37	pGEM-T-Easy	Amplified ORF (HEL37)	Blunt insert	Cloning
pGEM-T-Easy:Xyl-T36	pGEM-T-Easy	A. *thaliana* Xyl-T Targeting peptide	Blunt insert	Fusion Protein construction
pGEM-T-Easy:Xyl-T36::HEL12	pGEM-T-Easy	Xyl-T Targetting peptide + ORF (HEL12)	*Kpn*I-*Bgl*I-*Sal*I	Fusion Protein construct
pGEM-T-Easy:Xyl-T36::HEL37	pGEM-T-Easy	Xyl-T Targetting peptide + ORF (HEL12)	*Kpn*I-*Bgl*I-*Sal*I	Fusion Protein construct
pRSET C:HEL12	pRSET C	Excised ORF from pJET:HEL12	*Bgl*II	Protein expression in *E.coli*
pRSET C:HEL37	pRSET C	Excised ORF from pJET:HEL37	*Bgl*II	Protein expression in *E.coli*
pCambia1300:Xyl-T36::HEL12	pCambia1300	Xyl-T Targetting peptide + ORF (HEL12)	*Kpn*I-*Sal*I	Plant expression vector
pCambia1300:Xyl-T36::HEL37	pCambia1301	Xyl-T Targetting peptide + ORF (HEL12)	*Kpn*I-*Sal*I	Plant expression vector